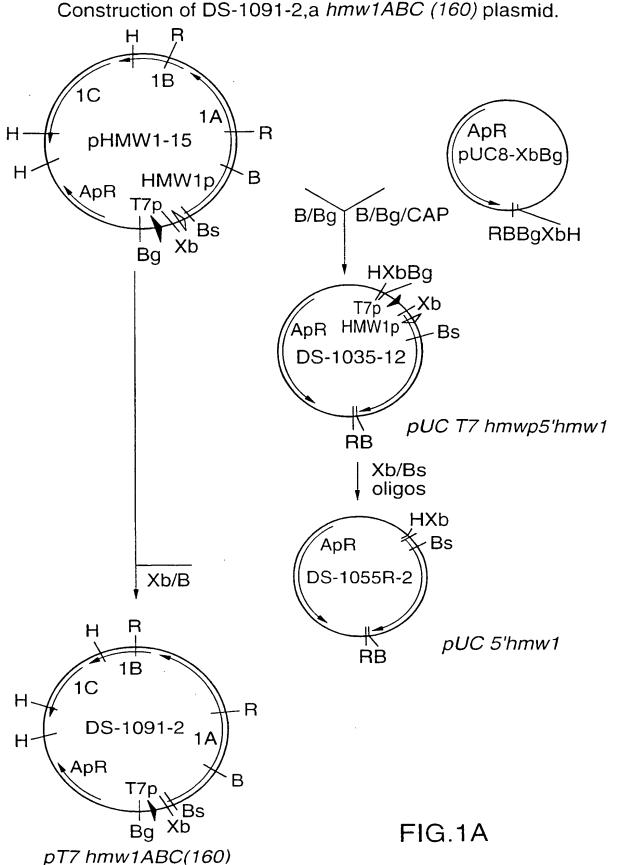
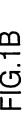
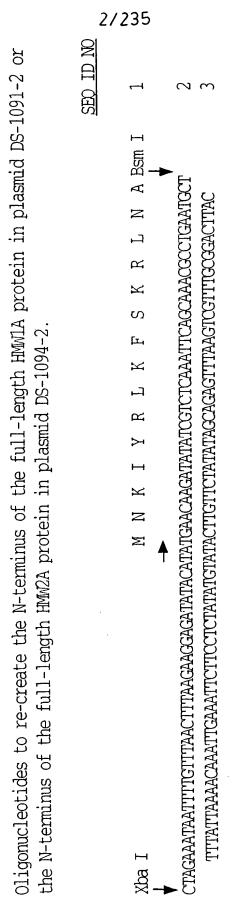
1/ 235



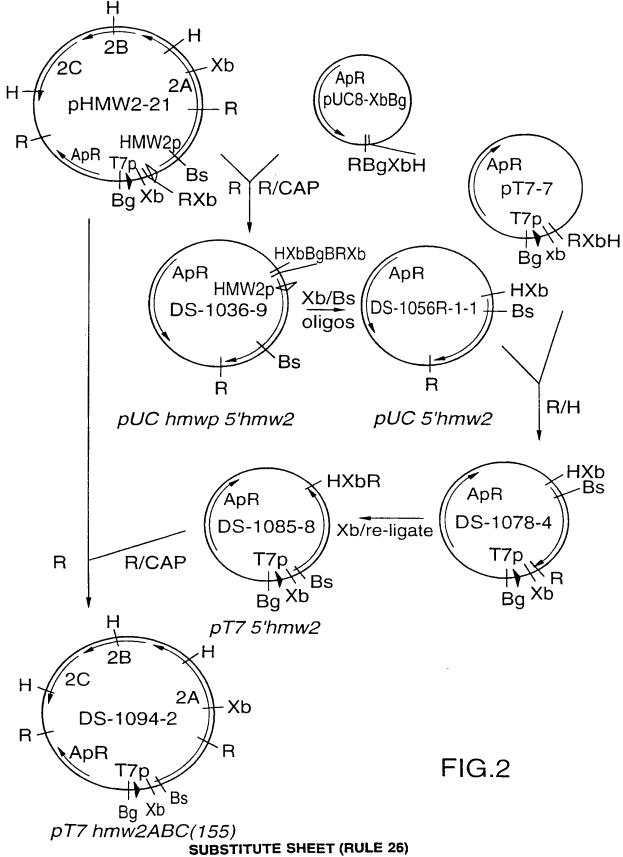
SUBSTITUTE SHEET (RULE 26)
BEST AVAILABLE COPY





3/235

Construction of DS-1094-2,a T7 hmw2ABC (155) plasmid.



Construction of DS-1046-1-1,a T7 hmw1ABC (125) plasmid.

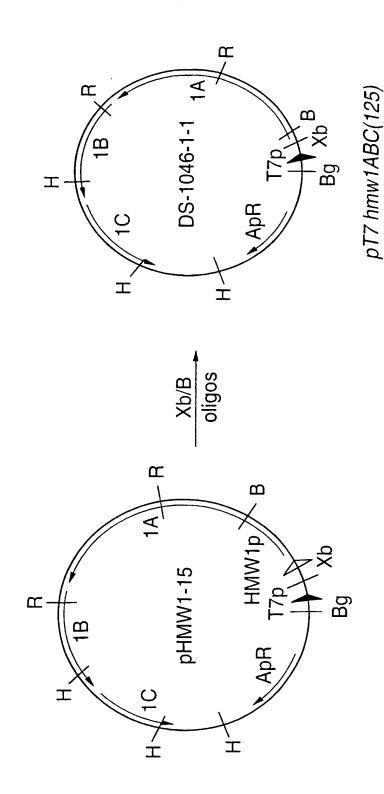
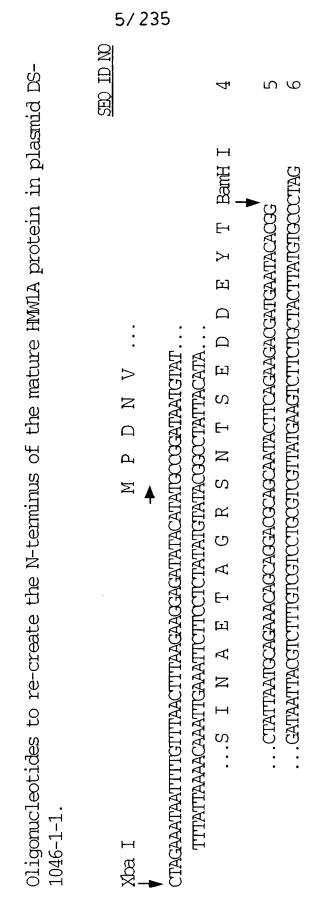
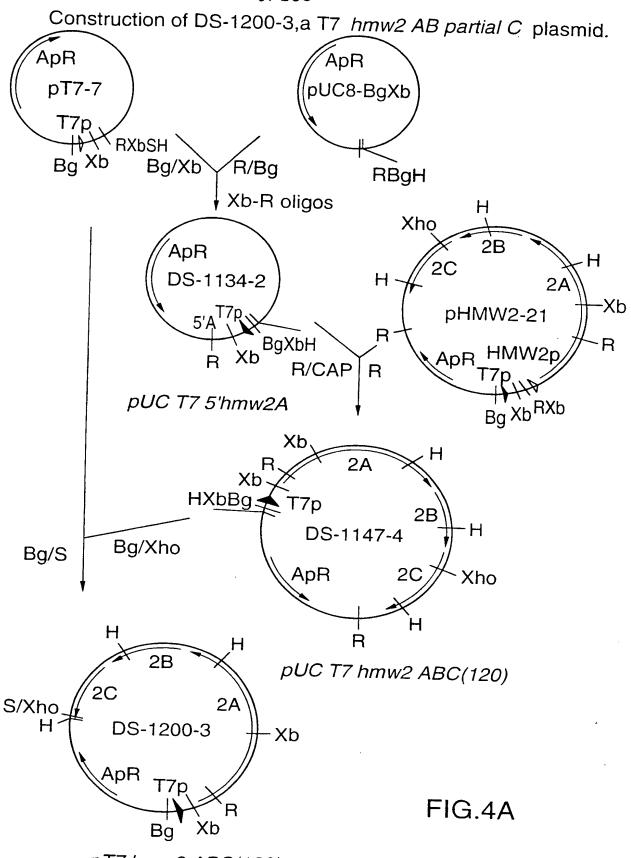


FIG.3A





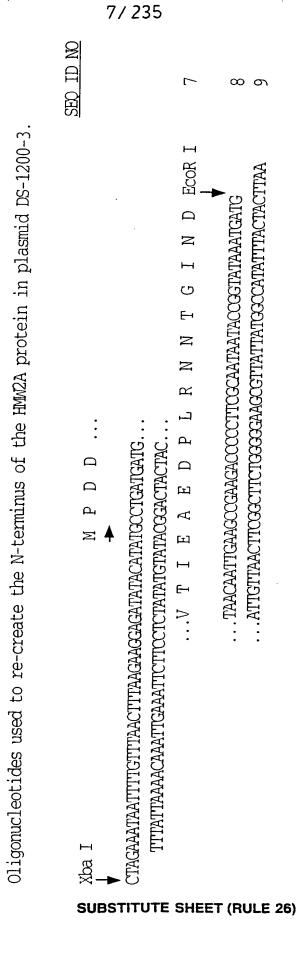
SUBSTITUTE SHEET (RULE 26)

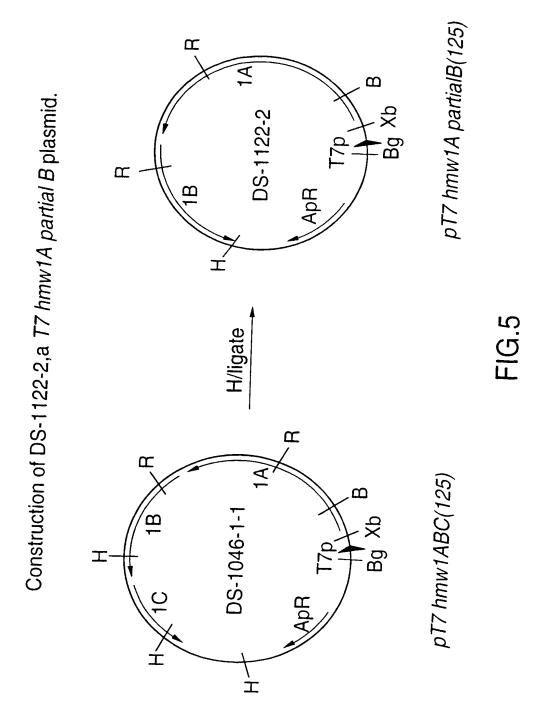


pT7 hmw2 ABC(120)

SUBSTITUTE SHEET (RULE 26)

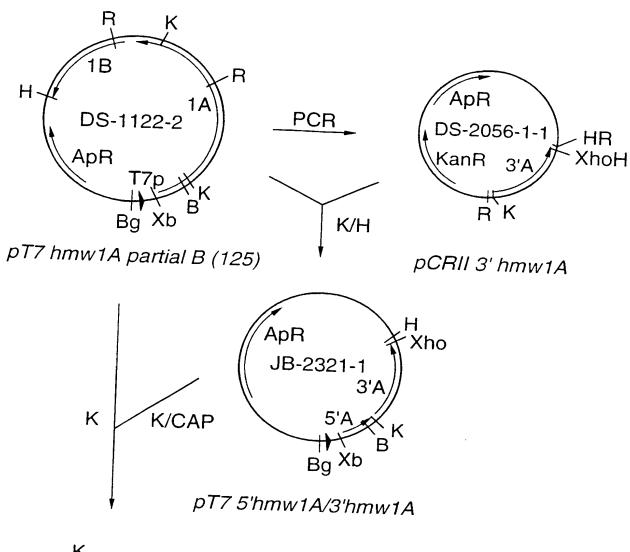
FIG.4B





SUBSTITUTE SHEET (RULE 26)

9/235 Construction of JB-2330-7,a T7 hmw1A (125) plasmid.



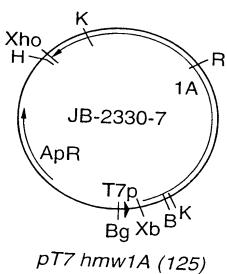


FIG.6A

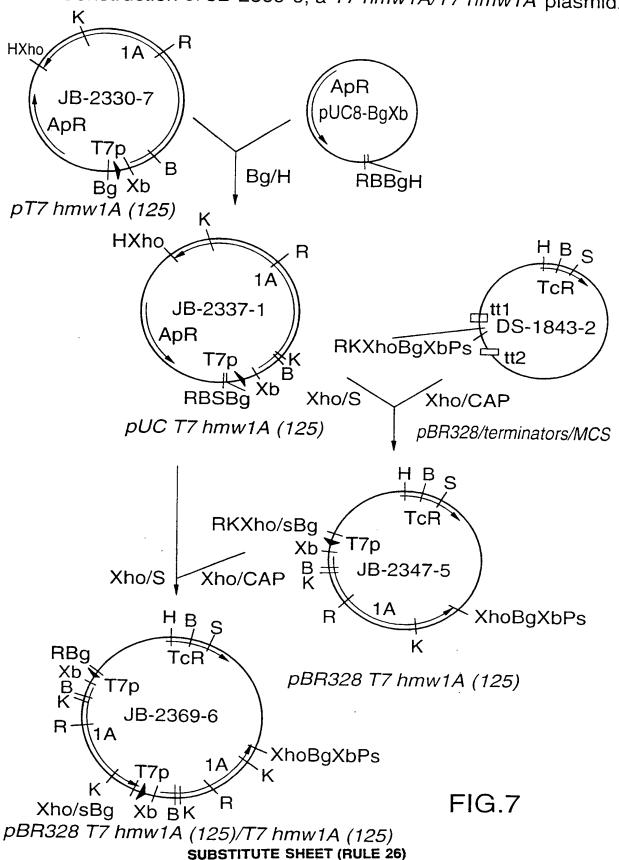
FIG.6B

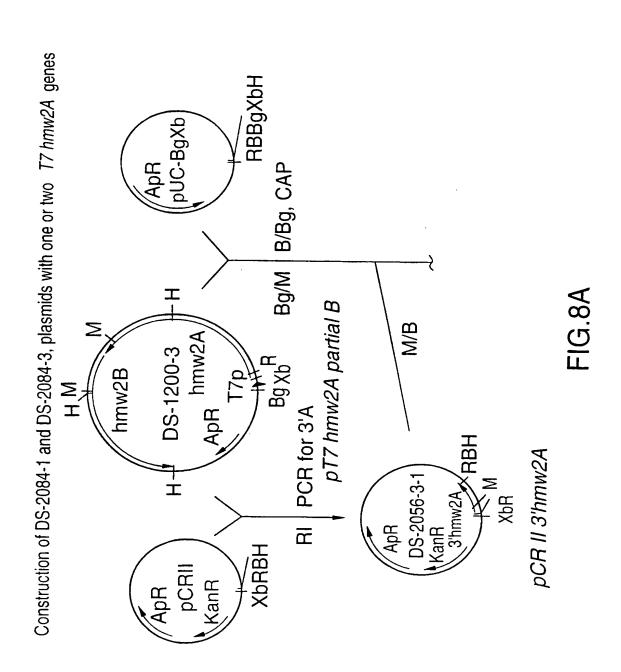
Oligonucleotides used to PCR amplify the 3'-end of hmwla.

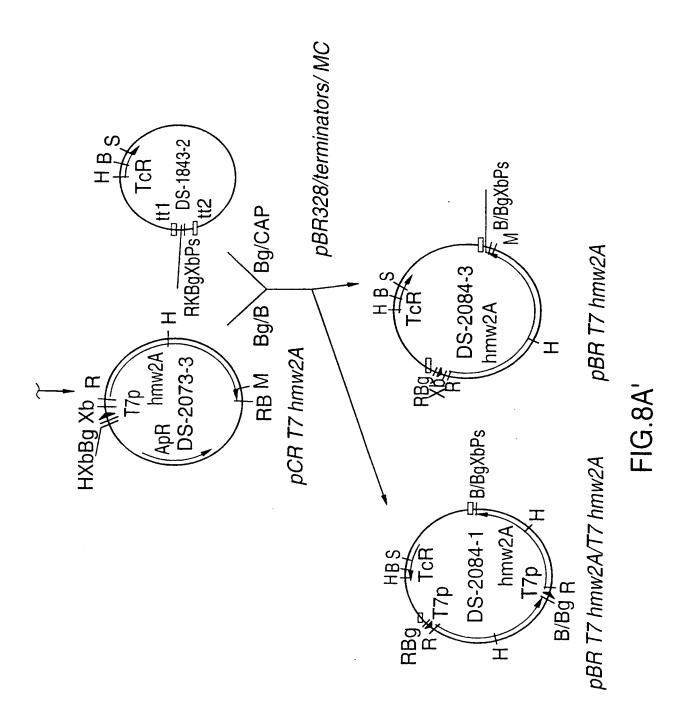
10 12 13 14 5358.JS 5359.JS 5, Xho I Hind III ø \Box × Ø œ ablaG z Ω K ⊱ G Z ഗ antisense sense 5, $\tilde{\sim}$

11/235

Construction of JB-2369-6, a T7 hmw1A/T7 hmw1A plasmid.







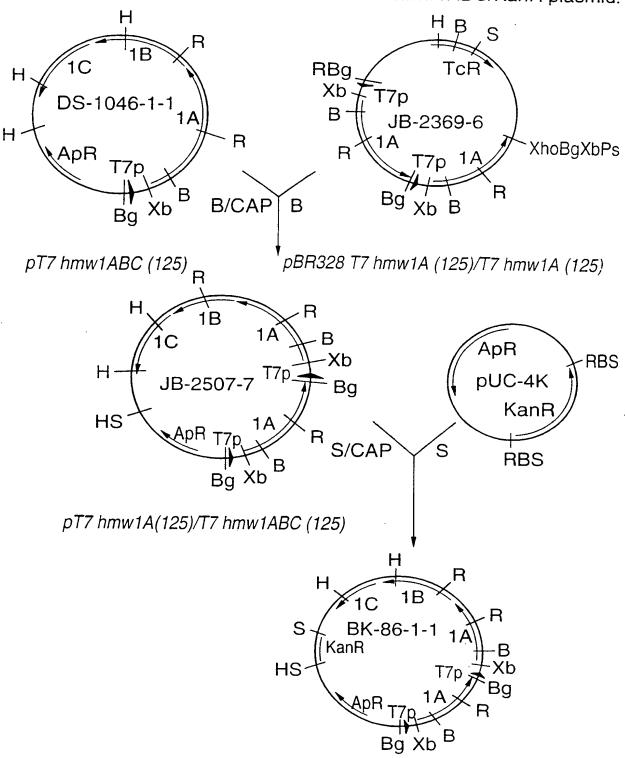
SUBSTITUTE SHEET (RULE 26)

14/235

SEO ID NO 15 17 18 19 5360.DC Oligonucleotides used to PCR amplify the 3'-end of hmw2A. 2 $\tilde{\mathcal{E}}$ BamH I > × 口 П G Ω 24 Д Ø EcoR I z FIG.8B antisense sense 2, ~

SUBSTITUTE SHEET (RULE 26)

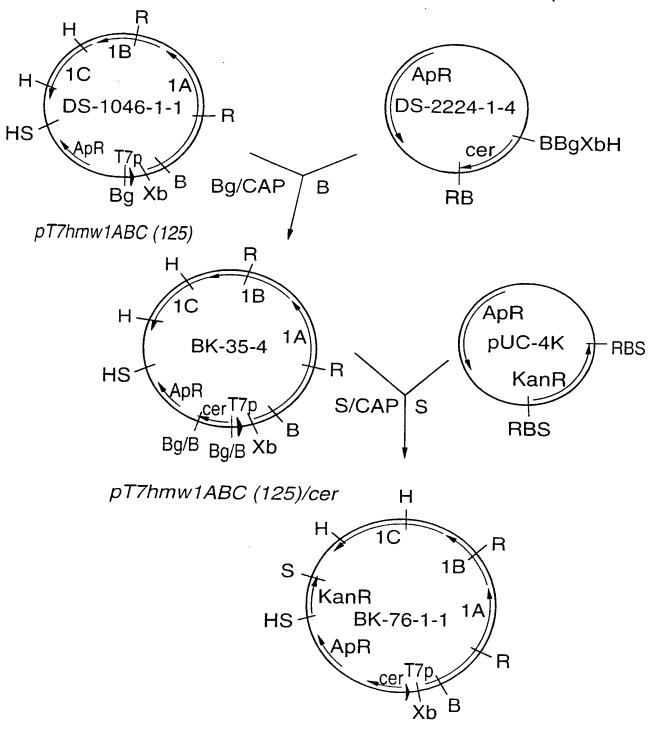
Construction of BK-86-1-1, a T7 hmw1A/T7 hmw1ABC/KanR plasmid.



pT7 hmw1A(125)/T7 hmw1ABC (125)/KanR

FIG.9
SUBSTITUTE SHEET (RULE 26)

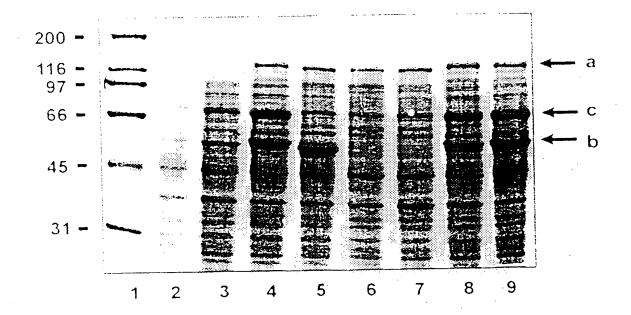
16/235 Construction of BK-76-1-1, a *T7 hmw1ABC/cer/KanR* plasmid.



pT7hmw1ABC (125)/cer/KanR

FIG.10

FIG. 11.



18/235 Purification of rHMW1 from E.coli

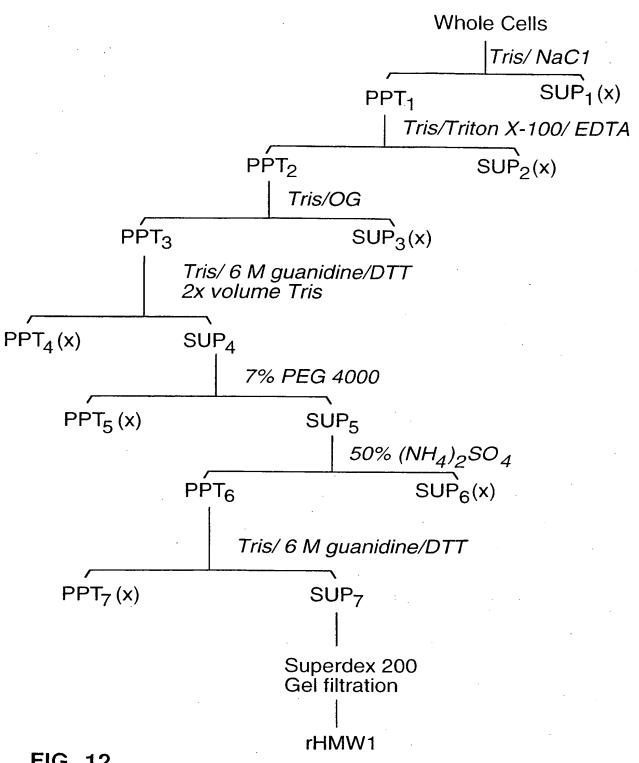


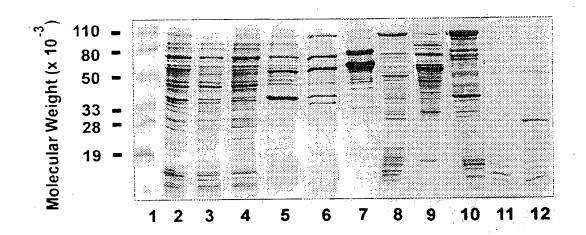
FIG. 12

All centrifugation is at 20,000g,30min. (x): to discard

19/235

FIG.13.

Extraction of rHMW1 from E. coli



20/235

Purification of rHMWs from E. coli

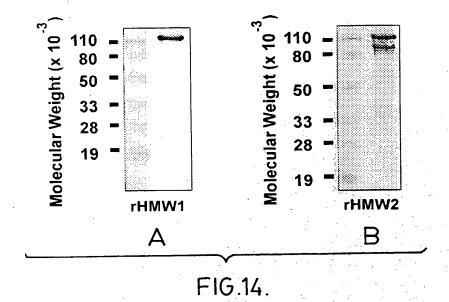
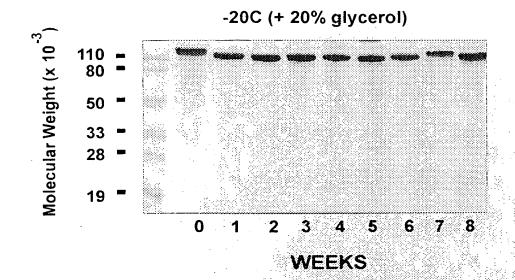


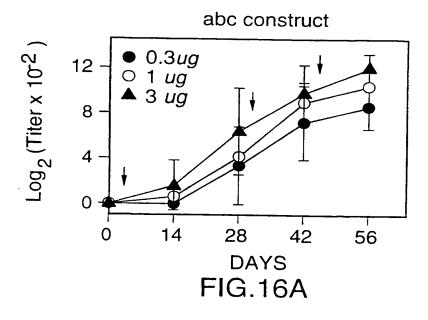
FIG.15.

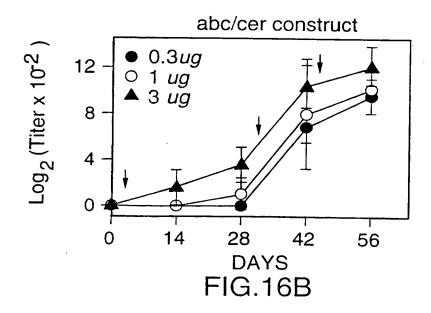
The Stability of rHMW1 (abc/cer)



22/235

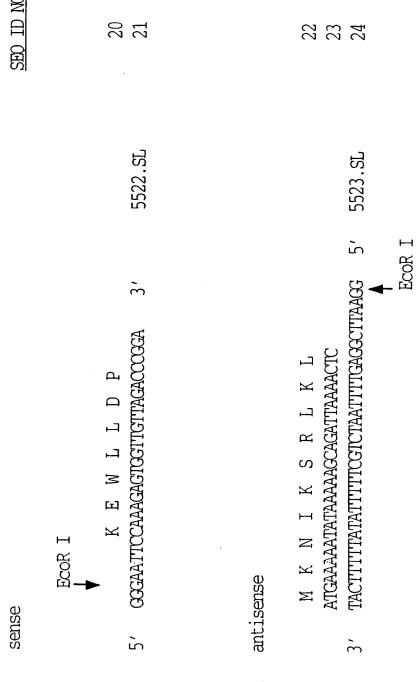
Kinetics of Antibody Response to rHMW1 in Mice





23/235

Oligonucleotides used to PCR amplify hmwA genes from NTHi strains.



SUBSTITUTE SHEET (RULE 26)

A A A 120

A A C (1110

A T

G \mathcal{O} Ø Ø Ø

G

 \mathfrak{O}

ල

Ø

G

K

G

A G 100

AA

PR0

ASP

.CGTAAC

ø

G A

ACA

	G G C		LYS
	GGACAATGTAT GGACAATGTAT 30SER ILE ASN ALA GLYCCATTAACGCAGGC	: : :	GLY ALA GLY GLY
	VAL GTA 30 ASN A A C (THR ACC 90	GLY
	ASN A A T ILE A T T	PRO CCT	ALA
	ASP GAC SER . CC1	PHE T T C	GLY
-	C C G 20	ASP G A T 80	:
Ice	LYS GLU TRP LEU LEU ASP PRO ASP ASN VAL A A A G A A T G T T A G A C C C G G A C A T G T A T 20 30 10SER ILE ASN ALACCATTAACGCA	ASP ALA SER PRO THR GLU ASP PHE PRO THR GACGCTTCACCAACAGAAGATTTCCCTACG 70 80 80	
sequer	LEU T T A	THR ACA	
umv1A	P LEU ; G T T G 10	PRO . C C A 70	
Joyc 1	TRP T G G	SER T C A	
18A	G A A	ALA G C T	
FIG.18A Joyc hmwlA sequence	LYS A A A	ASP G A C	
	su	JBSTITUTE	SH

	ASIN	AC	180	
	GLY	CAACTCTTGAGCGTATATTAAGTGGCAAC		
	SER	G T G		
	EE	I T A A	170	
	HE	ATA		
	ARG	CGT		
150	GLU	3 A G (:
	LEU	T T	160	ARG
	開	CTC		LYS
	TH	.CAA		ARG
140	:	:	:	ALA
				置
130				IIE
				ASN
				VAL
				r+1

A...

 \mathcal{O}

Ø

Ø

Ø

ΑT

K

 \vdash

ø

Ø G

Ø Ø

Ø

 \mathcal{C}

Ø

Ø

ں

K

G A 210 K Z, Ø Ø ø S Ø G C 200 ₽ C Ø €→ K ⊱ K T A 190 Ę \mathfrak{O} ⊱ Ø

_	Υ)	
<u> </u>	X)	
' (r	5	
L	Ī	_	

	25/23	5	
LYS A A A 240	LYS A A A 300	ASP G A T 360	ASP G A T 420
ILE THR VAL ASN SER ASP ILE ASN ILE LYS TCACAGTTAATAGTGATATCAACATCAAA 220	SER GLU 3 A G C G A A A 270 ASP ASN SER GLY VAL ASP ILE LYS G A T A A C A G C G G C G T G A T A T T A A A 280 290	TRP ILE	TU ASN SER GLY TAATTCAGGGC 390LEU LEU ASN ILLE THR THR LYS GLN GLY ASPTCTTAAACATTACAACTAAACAAGGAGAT TCTTAAACATTACAACTAAACAAGGAGAT
ILE ASN TCAACA 30	ASP GATA	TRP T G G	GIN C A A (
ILE A T C 230	VAL G T T C 290	GLY G G C T 350	LYS A A A C 410
ASP GATA 2	GLY GGCC	SER A G C	THR ACTA
SER AGTG	SER AGC	R GLY GLY SER TGGCGGAAGCT 330 LEU THR ILE TYR SER SER GLY TRP ILE TAACTATTTACTCCAGCGGCTGGATT 340 350	THR ACA
VAL ASN STTAATA 220	ACTCTGGAGCGAAA 260 270 ASN ASP ASN SERASN ASP ASN SERATGATAACAGCA	GLY ASN ILE THR SER THR THR GLY GLY SER GGCAATATCACTTCTACTACTGGCGGAAGCT 310 330 LEU THR ILE TYRTAACTATTAC 340	G C 0 ILE A T T
VAL AGTI	CGAA 270 ASN ASN ASN ASN 280	A A G C 330 ILE A T T T 340	LEU ASN SER GLY C T T A A T T C A G G G C 380LEU LEU ASN ILE T C T T A A A C A T T
THR	TRP SER GGAGC SN ASP ATGAT	CGGA THR	TTCA LEU
ILE T C A	LEU TRP T C T G G 260 ASN A T C	THR GLY GLY C T G G C G G A 20LEU THRT A A C T	LEU ASN SER TTAATTCA 80LEU LEUTCTTA?
	A C T 260 260	THR TAC 320	C 7 380
	ILE A A T A	THR	THR FAC
	LEU I C T A A	SER LTCTA	ILE SATTA
	HIS CCATC 250	E THR CAC 310	S ASN A A A C 370
	SER T C C C 250	ILE ATC	LYS A A A A
	ASP SER SER HIS LEU ILE GACAGCTCCCATCTAAT 250	GLY ASN ILE THR GCAATATCACTT 310	ILE HIS LYS ASN ILE THR A T T C A T A A A A C A T T A C G 370
	ASP G A C	G G C	ILE A T T

FIG. 180

ASN A A C

ASN A A T

666

AAA

GAA

	26/235		
GLY G G C 480	LEU C T T 540	THR ACT 600	
ALA 3 C A C	LEU	GLY 3 G A 2	
THR	GLY G G 7	ASN GLY	
ILE A T T A 470	THR GLY ACAGGG1 530	PHE TTA 5590	
IHR CTA 4	GLY 3 G A A 5	PHE 55	
JLY ' G G A	ILE (ASN I	
 GIN GLY THR SAAGGGACTA)	 T GLY J	 T ILE P	
	PHE GLU ASN ALA TTTGAAAACGCCT 500 510 SER LEU ASN GLY ILE GLY THR GLY LEU LEU CCCTAAACGGTATTGGAACAGGGTTACTT 520 540	U GLY ASN ASN AGGAAATT 570PHE GIN ILE ILE ASN PHE PHE ASN GLY THRTCCAAATCATAAACTTTTTAACGGAACT	SER
450 THR GLY ACAGGT 46	ASN ACG EU A	LEU GLY ASN ASN TAGGAAATAAT 60PHE GLN ILETCCAAATC 58	ILE TCT
) ILE T T T A	E GLU ASN TGAAAACSER LEU	U GLY ASN AGGAAAT PHE GLN TCCAAA	ASN A C A
440I	PHE GLU ASN TTTGAAAAC 500SER LEU CCCTA.	LEU (17 A G 560Pl	VAL ASN ILE GTAAACATCI 620
4		ASP 5 A T T 5	LYS 1 1 A A G 62
	PHE TA	G G G	JLY 1 GGA
	J.Y. J. G. T. T. G. T. T.	LYS ARG AAAGGO	SER GLY
430	LYS GLY A A A G G T 7 490	ILE I T C A 550	LE S TTT 610
	ASN GLY LYS GLY PHE ARG A A T G G T A A A G G T T T T A G G 490	PHE ASN ILE LYS ARG ASP LEU GLY ASN ASN TTTAACATCAAAAGGGATTTTAGGAAATAATT 550 570PHE GLN ILE ILETCCAAATCATA 580	LEU ASN ILE SER GLY LYS VAL ASN ILE SER TTAAATATTCAGGGAAAGTAAACATCTCAA 610 620
	ASN GLY .ATGGT <i>P</i>	HE ATTA	EU A TAA
	A	H	H

SUBSTITUTE SHEET (RULE 26)

(7	
C	α	
٦	_	
(ワ	
ī	L	

	27/235		
SER AGT 660	LYS A A G 720	TYR T A T 780	ALA 3 C A 840
MET VAL ILE PRO LYS LYS TRP ASP TYR SER TGGTCATACCTAAAAAATGGGATTATAGT 640 650	R TRP ASN VAL TTGGAACGTAA 690THR HIS LEU ASN VAL SER GLU GLY SER LYS CCCATTTAAATGTTTCGGAAGGCAGTAAG	R ARG GLY ASP CAGAGGATG 750ASP THR ALA GLY THR LEU ASN THR PRO TYR A CACTGCAGCACCCTTAACACCCCTTAT 780	E ASN LYS ASP CAACAAGACA 810 THR ILE PHE ASP VAL LYS GIN ASN GLY ALA CTATCTTTGATGTTAAACAAACGGGGCA 830 840
ASP GATT	GLY G G C A	THR	ASN A A C G
TRP T G G G 650	GLU G A A C 710	ASN A A C A 770	LYS GLN AACAA? 830
LYS A A A	SER TCG	LEUCTTA	LYS A A A
LYS	VAL	THR	VAL G T T
ILE PRO TACCTA 640	TYR TRP ASN VAL TATTGGAACGTAA 680 690 THR HIS LEU ASN CCCATTTAAATC	ASP G A T G 750 ALA GLY C A G G C P	PHE ASN LYS ASP TTCAACAAAGACA 800 810THR ILE PHE ASPCTATCTTTGATG
ILE ATA:	CGTA 7 690 LEU 7 1 1 A A A 700	ASP A G A T 750 ALA G C A G	ASP A G A C 810 PHE T T T G
VAL GTCA	GAAC GAAC HIS	SER ARG GLY ASP C T C C A G A G G A G A T G 740ASP THR ALA GLYA C A C T G C A G G C	E ASN LYS CAACAAATHR ILE CTATCT
NET T G (TYR TRP ATTGG 1 80THR H CCC	SER ARG C C A G A 40ASP A C A	ASN CAAC THR CT
• • •	TY 2 T A 680	SE 740	PH T T 800
	\sim		ILE SER
	ARG	ILE	ILE
	G G G G C 670	U THR CAC1 730	N GLY C G G T 790
	ARG 2 A G G G 670	LEU	ASN A A C
	LYS PHE ARG GLY ARG THEAAAG GGGGGGAAC 670	PHE ASN LEU THR ILE ASP TTTAACCTCACTATCGA 730	ASN LEU ASN GLY ILE SER A A T T T A A A C G G T A T A T C A 790
	LYS A A A T	PHE TTT	ASN A A T

G A A 1050

CJ

AGA

ATTT. 1040

G A

T T A

CGAG'

G T

. G

G G A

LEU LEU ALA TTCTCGCC	ASN ALA SER ATGCTTCA 1020	
VAL SER GLY A G T T T C A G G A G 930 GLY ASN VAL THR PHE LYS LEU G A G G G A A T G T C A C T T T C A A A C T ' 940 950	PRO GLY VAL FCCGGTGTAT 990 PHE ILE ASN SER LYS HIS PHE ASN TTATAAATTCTAAACACTTTAA 1000 1010	PHE ARG THR GLU
SER 1 T C 7 920	1 THR 980 980	
ILE FAT	GIN C A 7	CIN
ASN ASN	ALA G C T	LEU
GLY GGA 10	THR ACC 70	SER
ASN A A T 9.	SER TCT 97	SER
PHE TTC	SER T C C	GLY
SER T C A	SER T C A	GLY
	AGTTTCAGGAG AGTTTCAGGAG 930GLY GLY ASN VAL THR PHE LYS LEUGAGGGAATGTCACTTTCAAACTTCGAGGGAATGTCACTTTCAAACTT	SER PHE ASN GLY ASN ILE SER VAL SER GLY 1 C A T C C A T G C A A T T T C A G T T T C A G A G 910 920 930 1 G A G G A A T G T C A C T T C A A C T C T C C G C C C C C C C C C C C C C

L	1	-
(X)
٦	_	_
(F	5
-	Ξ	_
L	1	_

	29/235		
ASN A A T 1080	11.E A T T 1140	ASN A A T 1200	ASN A A C 1260
ILE A T A 7	G G C	GLY G G C	ASN A A T
LEU TTG	GLU GAA	GLY G G A	ILE A T C A
C T T C 1070	GIN VAL GLU GLY CAAGTTGAAGGCA 1130	PHE ALA GLY GLY TTGCTGGAGGC 1190	ALA THR CTACTA 1250
GLY GGC 1	GIN C A A	PHE TTT	AIA G C T 2
THR LYS VAL GLY PHE LEU CAAAAGTCGGCTTCTTGA 1060 1070	GLY GGTA 1110 SER LEU LEU 'CGCTCTTGC	THR	ASN AAT(
LYS A A A 50	 LEU C T C	ALA G C T A 1170 ASN ILE ASN ILE A C A T A A 1180	GILY G G G C
GLY SER THR LYS VAL GLY PHE LEU ILE ASN GCTCAACAAAGTCGGCTTCTTGATAAAT 1060 1060 1070	N ALA THR GLY GLY TGCCACCGGAGGTA 1100 1110ASN ILE SER LEU LEU GLN VAL GLU GLY ILEACN ILE OCTCTTGCAAGTTGAAGGCATT 1140ACATATCGCTCTTGCAAGTTGAAGGCATT	S GLY VAL VAL ALA AGGCGTTGTAGCTA 1160LYS LYS ASN ILE THR PHE ALA GLY GLY ASNAAAAACATAACCTTTGCTGGAGGCAATAAAAAACATAACCTTTGCTGGAGGCAATAAAAAACATAACCTTTGCTGGAGGCAAT	ILE THR PHE GLY SER LYS ALA ILE THR A T C A C C T T T G C T C C A A A G C C A T A A C A G 1230 1.GLU ILE GLU GLY ASN ALA THR ILE ASN
SER TCAA	GLY C G G , ILE A T A	VAL IGTA LYS AAAA	ILE ILE ATC
GLY G C T	ALA THR GLY GLY CCACCGGAGG' 100 111ASN ILE SERACATATCG	GLY LYS GLY VAL VAL VAL GTA (GTAAGGCGTTGTA) 1160LYS LYS AAAAA	S ALA A G C CGLU A A A
• • •	AI.	GLY A G G G 1160	LYS 3 A A 7 1220
		LYS T A A J	ILE THR PHE GLY SER LYS TCACCTTTGGCTCCAAGA 1210
	LEU	GLY G G	SER T C (
	THR A C C	ILE ATTG 50	G G C C.
	ASN ASP LEU THR LEU AS A A T G A T T T A A C C C T A A A 1090	ASP GLY MET ILE GLY LY GACGGATGATTGGTAA 1150	PHE (T T T G 1210
	ASP G A T	G G G	THR A C C
	ASN A A T	ASP G A C	ILE A T C

T T T 1440

PCT/CA99/00938

	LEU TT	ILE A T (PÆ T T C
	PHE ASP ASN HIS GIN LYS PRO LEUTTGACAACCATCAAAAACCTTT? 1300	N SER GLY ASN A T A G C G G C A A C C 1350LEU THR ALA GLY GLY ASN VAL ILET T A C C G C T G G C G C A T G T T A T C 1360	N ASN GLY ALA CAATGGCGCCA 1410ASN LEU LYS ALA ILE THR ASN PHE
	LYS A A A	ASN A A T	THR ACA
	GLN C A A	G G C	ILE A T C
ASP 3 A T T 1290	HIS C A T 10	 GG G C	A ALA G C T
ASP ; G A 7 1290	ASN H. A A C C 7 1300	ASN 1350 1350 ALA G G C T G (ALA G C C 1410 LYS A A A G
SER TCG	ASP GAC	GLY GGC THR	GLY GGC LEU
GLY GGT	. T T .	SER AGC .LEU	ASN A A T ASN A T (
1 ILE TATC 1280	: : :	ASN T A A T 1340	ASN 2 A A C 1400
LEUCTT		ILE ATT	VAL GTT 1
THR ACT		ILE A T C	THR A C C
VAL GTC 70		VAL G T C	LEU CTT O
ASN V7 A A C G 1 1270		ASP V7 G A T G 1 1330	ASN L AATC' 1390
ASN ALA ASN VAL THR LEU ILE GLY SER ASP A A C G C T A A C G T C A C T C T T A T C G G T T C G G A T T 1270 1290		LYS LYS ASP VAL ILE ILE ASN SER GLY ASN A A A A A A G A T G T C A T C A T T A A T A G C G G C A A C C 1330 1340LEU THR ALA GLY 1360T T A C C G C T G G C	ASN GLY ASN LEU THR VAL ASN ASN GLY ALA ACGGAATCTTACCGTTAACAATGGCGCCA 1390 1410ASN LEU LYS ALA
ASN A A C		LYS A A A	ASN GLY ASN LEU THR VAL ASN ASN GLY ALA A A C G G A A A T C T T A C C G T T A A C A A T G G C G C C A 1390 1410 1ASN LEU LYS ALAASN LEU LYS ALA

30/235

AACTATT

ASN ILE SABTATA 1380

> ASN VAL GLY GLY LEU PHE ASP ASN LYS A A T G T A G C G C T T G T T G A C A A A A 1450 1450

SUBSTITUTE SHEET (RULE 26)

ATTCAAATATCTCCATTGCTAGGAGGG ATTCAAATATCTCCATTGCTAGGAGGGGGG	ALA LYS PHE LYS ASP ILE ASN AGN THR SER GCTAAATTTAAAGATATCAATAACACCAGTA 1510 1520 1530 SER LEU ASN ILE THR THR ASN SER ASP THRGCTTAAATATTACCACCAACTCGACACC 1560	THR TYR ARG THR ILE ILE GLU GLY ASN ILE ACTTACCGTACCATTATAGAAGGTAATATAA 1590 1570THR ASN LYS ALA GLY ASP LEU ASN ILE ILECCAACAAGGTGATTTGAATATCATT 1620	ASP ASN LYS GLY ASN ALA GLU ILE GIN ILE GATAATAAAGGTAACGCTGAAATCCAAATTG 1630 1640 1650 CLY GLY ASN ILE SER GLN LYS GLU GLY ASN GCGGCAATATCTCGCAAAAGAAGGTAAT 1680 1670 1680
	ALP	THE	ASF
	G C	A C	G A

LEU THR ILLE SER ASP LYS ILLE ASN ILLE CTCACGATTCGATAAATTAATCA 1710 1690THR ASN GLN ILE THR ILLE LYS LYS GLY VAL CTAACCAGATAACAATCAAGAGGGTGTT 1740	ASN LYS GLU ASP SER ASP SER THR ALA A A T A A A G C A C G C C A A 1770 1750ASN ASN ALA ASN LEU THR ILE LYS THR LYSASN ASN ALA ASN LEU THR ILE LYS THR LYSAC A A T G C T A A T C T A A C C A T T A A A A C C A A A 1800	GLU LEU GIN LEU THR GLY ASP LEU ASN ILE G A A T T G C A A T T A A C G G G A C C T A A A T A T T T 1810 1820 SER GLY PHE ASP LYS ALA GLU ILE THR ALA C A G G C T T C G A T A A A G C A G A A T C A C A C C 1860	LYS GLU GLY ALA ASP LEU ILE ILE GLY ASN A A A G A G G G C C G A T T T A A T C G G T A A T A 1870 1880
--	--	--	---

LYS A A A 1920	THR A C A 1980	ASP G A T 2040	HIS C A C 2100
SER ASP ASN ASN ASN ALA ASN ALA LYS GTGATAATAACAACAATGCTAATGCTAAA 1900	L LYS ASP SER TAAAGATTCGA 1950 LYS ILE SER ALA GLY SER HIS ASN VAL THR AAATCTCTGCTGGCAGTCACAATGTAACA AAATCTCTGTTGTTGTTACA	R SER ASN GLY CTCTAATGGCA 2010 ASN ASN ASP ALA GLU SER ASN ASN GLY ASPASN ASN ASP ALA GLU SER ASN ASN GLY ASPASN ASN ASP ALA GLU SER ASN ASN ASN ASPASN ASN ASP ALA GLU SER ASN ASN ASN ASP	N ALA LYS ASN TGCAAAAATG 2070VAL THR VAL ASN ASN ILE THR SER HISTAACAGTAACAACAATATTACTTCTCACTAACAGTAACAACAATATTACTTCTCAC
ASN A A T G	ASN A A T	ASN ASN GLY AATAATGGC 330	THRACTI
ALA : G C T A 1910	HIS CACA	ASN 2 A A T 2030	ILE 2090
ASN A A T G	SER AGTC	SER A G C A 200	ASN A A T A 20
ASN A A C A	G G C	GLU GAA	ASN A A C
ASN A A C A	3 A 3 A ALA G C T (2 A 3 ALA G C T (I G D ASN A A C A
ASN <i>P</i> AATA 1900	GIN VAL LYS ASP SER C A G G T T A A A G A T T C G A 1940 LYS ILE SER ALA A A A T C T C T G C T	GLU THR SER ASN GLY GAAACCTCTAATGGCA 2000 2010 ASN ASN ASP ALAATAATGACGCT	ILE ASN ALA LYS ASN A T T A A T G C A A A A A A T G 2060VAL THR VAL ASN T A A C A G T A A A C
SER ASP GTGATA	L LYS ASP TAAAGAT LYS ILE AAATC	THR SER ASN CCTCTAAT 00ASN ASNATAATG	ASN ALA LYS ATGCAAAAA 60VAL THR VTAACAG
G T	LYS T A A A G LYS I A A A	SER C T C ' ASN	N ALA TGCA VAL
• • •	VAL 3 G T T 7 1940]	THR A C (2000	ASN 1 A A ' 2060
	ASN A A C	VAL G T P	LEU THR
	R PHE: CITI	R LYS IT A A A G 1990	R LEU ; C T T A 2050
	THR ACC 19	ASN SER 1990	SER A G C 20
	LYS VAL THR PHE ASN A A G T A A C C T T T A A C 1930	LEU ASN SER LYS VAL TAAACAGTAAAGTA 1990	SER THR SER LEU THR GCACCAGCTTAACT 2050
	LYS 1 A A	LEU	SER A G C

ACAGTGGCAGCAGCC

E E

AA

CAA

GGT

ACA

ASIN

		ASIN	AAT	2160				CIU	GAA	2220					ILE	A T C	
		VAL THR THR LYS ALA GLY THR THR ILE ASN	ТТАССАССАААВССВСАСАСАССАТТААТ					LYS THR GLY ASP ILE LYS GLY GLY ILE GLU	A A A C A G G T G A T A T T A A A G G T G G A A T T G A A						SER GLY ASP THR LEU ASN VAL SER ASN ILE	G C G G C G A C A C G C T T A A T G T A A G T A A C A T C	
		訊	ACC					GLY.	G G A						SER	AGT	
		THY	ACA	2150				GLY	G G T	2210					VAL	GTA	
		GLY	ე ე	(7)				LYS	AAAA	2					ASN	AAT	C
		ALA	ე ე					II E	АТТ						LEU	СТТ	
•	G	LYS	AAA	0	:	A	:	ASP	G A T	00		:	3 A	:	黑	ACG	c
ASN	A A T (2130	訊	ACC	2140	ALA	0001	2190	GLY	G G T	2200		ALA	0001	2250	ASP	GAC	
THR ALA SER GLU ASN	ACTGCGTCAGAAAATG 2120 2130	開	ACC		VAL GLU VAL THR ALA	GTAGAAGTAACAGCCA		開	ACA			VAL ASN ILE THR ALA	GTAAATATTACAGCGA		GLY	C C C	
SER	TCA	.VAL	₽		VAL	GTA		TIYS.	. A A	•		ILE I	ATT		SER.	ე ე	
ALA	2120 C G	:	:	:	CIU	GAA	2180	:	:	•		ASIN	AAT	2240	:	:	
景					VAL							VAL					
TIE	ATC				SER	AGC						ASIN	AAT				
LYS THR VAL ASN ILE	AAAACAGTAAATATC 2110				GLY	GCAACCATAGGTAGC	70					CLY	TCCAATTCCGGTAAT	30			
VAL	G T A A 2110				ILE	ATA	2170					SER	$T \subset C$	2230			
出	ACA				出	ACC						ASN	AAT				
LYS	AAA				ALA	GCA						SER	$T \subset C$				
				su	BST	ITL	JTE	SH	EE.	T (R	ULI	E 26	5)				

c	X)
۲ (r	- 5
Ĺ	ī	_

	35/235		
SER T C A 2340	GLY G G C 2400	VAL G T A 2460	THR A C A 2520
SER GLY ALA VAL THR THR THR LYS GLY SERCAGGTGCCGTAACAACCACAAAGGATCA 2320 2340	Y ASN ALA ASN TAATGCAAATA 2370ILE THR THR LYS THR GLY GLU ILE ASN GLYTTACAACCAAACAGGTGAAATTAATGGC	Y ASN VAL ASN TAATGTAAATA 2430 ILE THR ALA SER GLY ASN THR LEU ASN VAL TTACAGCGAGCGAATACACTTAATGTA TTACAGCGAGCGAATACACTTAATGTA	N VAL THR VAL TGTAACAGTAA 2490 THR ALA ASN SER GLY ALA ILE THR THR THR CAGCAAACTCAGGTGCCATAACAACCACA CAGCAAACTCAGGTGCCATAACAACCACA
THR LYS GLY CAAAAGGA 30	ILE A T T A	LEU	THR ACAP
C A C A 2330	GLU F G A A A 2390	THR 1 A C A C 2450	ILE SATAA 2510
THR ACCA 23	GLY G G T	ASN A A T A 24	ALA G C C A 25
THR ACAA	THR A C A	G G C	G G T C
VAL GTA? 20	ASN A A T A 2370 HR LYS C C A A A A 2380	ASN A A T A 2430 ALA SER C G A G C (240	VAL G T A A 2490 ASN SER A C T C A C
ALA V G C C G 2320	ASN A A A T I 2370 THR I A C C A 2380	ASN A A A T 7 2430 ALA 8 G C G A 2440	VAL 4 G T A 2 2490 ASN S A A C T A A C T
G G T	TGCA THR ACA	VAL TGTA THR ACAG	THR AACA ALA GCA
SER C A (T G G T A A T G C A A A T A 2360 2370 ILE THR THR LYS T T A C A A C C A A A	C G G T A A T G T A A A T A 2420 2.11E THR ALA SER T T A C A G C G A G C	A A A T G T A A C A G T A A 2480 THR ALA ASN SER C A G C A A A C T C A 2500
	GE G G 2360	GL C G G 2420	AS A A A 2480
			•
	THR	ALA	GLY G G C
	SN ALA A T G C <i>I</i> 2350	SER TCAG	THR A C 1
	ASN A A T G 2350	VAL LYS STTAAAT 2410	ILE TA TCA 2470
	THR ILE ASN ALA THR THA ACTATTAATGCAACAAC 2350		SER ASN ILE THR GLY GLI AGTAACATCACTGGTCA 2470
	THR ACTA	GLU GAAG	SER A G T A

G G A A... 2730 ...

 \mathcal{O}

T A A

GAT

T A A (2720

 Γ

AAG

 \circ

A C

A C

AC

A A A 2710

GACAACAGGTG GACAACAGGTG 2550ASP ALA ASN ILE THR THR GIN THR GLY ASN A T G C A A A T A T T A C A A C C C A A C G G T A A T A T G C A A A T A T T A C A A C C C A A C G G T A A T A T G C A A A T A T T A C A A C C C A A C A G G T A A T A T G C A A A T A T T A C A A C C C A A A C A G G T A A T 2560 2570	R SER SER GLY CAGTICTGGTT 2610SER VAL THR LEU ILE ALA THR GLY GLN THR CTGTGACGTTATTGCAACTGGACAAACT C7676 2620	R GLY ASP THR AGGTGACACTG 2670VAL THR ILE THR ALA ASP LYS GLY LYS LEUTACCATTACTGCGGATAAAGGTAAATTA 2700	
AAC	T G C	ASP G G A	
THR FACAA	ILE LATT C	ALA	
GLY G G T G 2550 ASN ILE A T A T T A 2560	GLY G G T T 2610 THR LEU C G C T T A	 F.G THR A.C.T	:
GLY G G T (2550 ASN 1 A A T A 2560	GLY G G T T 2610 THR I A C G C 2620	THR A C T (2670 LLE A T T A 2680	GLY
THR ACA ALA GCAA	R SER SER CAGTTCTSER VALCTGTGA	SER GLY ASP CAGGTGACA 60VAL THR 1TTACCA	ASN
A THR GACAASP A T G	SER AGTT SER V.	GLY GGT VAL	ILE
ASN ALA THR THR GLY A A C G C G A C A A C A G G T G 2550 ASP ALA ASN ILE A T G C A A A T A T T	SE C C	됐 20 05	LYS
ASN A A C	G A A T 26	ILE ATTT 260	SER
ILE A T T	VAL G T T	ASN A A T	SES
THR ACTA 30	LYS A A A G 30	GLY GGTP	開
SER 1 T C A A 2530	GLY I G G T A 2590	VAL (GTAGT)	GLN
GLU GLY SER THR ILE GAAGGCTCAACTATT 2530	ILE ASN GLY LYS VAL ATTAATGGTAAAGTT 2590	LEU ALA VAL GLY ASN CTTGCTGTAGGTAAT 2650	訊
GLU GAAG	ILE A T T A	LEUCTT	景

_	_	_
C	X	5
7		_
(Т	;
_	=	_
L	1	_

THE LYS SER VAL THR THR SER SER GIN C T A G G G T G T A A C C A C C T C A G C C A A T 2740 2740 R ILE SER GLY ASN 2780 C G G T A A T A 2780 C G G T A A C C G T C G A C C G G T A G C C C G G T A A G C G T T A G T G C G A C C G G T A G C C G G T A A G C G T T A G T G C G G T A G C C C G G T A A G C G T T A G T G C G G T A G C C C G G T A G C A A 2840 2840 A A A A A T T G A A G C A A 2850 A A A A A T T G A G C A A 2860 A A A A A T C T C G G T G A G C T A A T G T A A C A G C C A A A C A G G T G A G C T A A T G T A A C A G C C A A A C A G G T G A G C T A A T G T A A C A G C C A A A C A G G T G A G C T A A T G T A A C A G C C A A A C A G G T A A T G T T A C A G C A A T A C T G A T A C A A T C T C T G C C A A T A C T G A T A C A G T A A A T G T T A C A G C A A T A C T G A T A C A G T A A A T G T T A C A G C A A T A C T G A T A C A G T A A A T G T T A C A G C A A T A C T G A T A C A G T A A A T G T T A C A G C A A T A C T G A T A C A G T A A T G T T A C A G C A A T A C T G A T A C A G T A A T G T T A C A G C A A A T A C T G A T A C A G T A A T G T T A C A G C A A T A C T G A T A C A G T A A T G T T A C A G C A A T A C T G A T A C A G T A A T G T T A C A G C A A T A C T G A T A C A G T A A T G T T A C A G C A A T A C T G A T A C A G T A A T G T T A C A G C A A T A C T G		37/235		
R III 2780 2780 2840 2840 2900 2900	SER SER GLN SER CTCAAGCCAATCA 2750 . 2760		N VAL THR SER ALA TGTAACAAGCGCA 2870	ALA ASN THR ASP GCAAATACTGAT 2930
GLY G G T A C C A C A	THR LYS SER VAL THR THCTAAGAGTGTAACCA(ER GLY THR ILL GTGGCACAAT 2780	LA GLY SER LY CAGGCTCAAA 2840	GLY GLY TH GCGGTAC 2900

ARG A G A 7

GCA

LYS ASP GLY A A A G A T G G C

ATT

LEU TTA

AA

THR A C T

(C	
C	Q	
T	_	
(5	
Ī	工	

	38/235		
THR A C C 3000	GIN C A G 3060	THR A.C A 3120	
LEU TTA	GLY G G T	VAL G T G	
THRACT	ASN A A T	ASN A A T	
VAL r g T G 2990	SER A A G C 3050	c G C C 3110	
ALA G C T	SER TCA	ALA G C C	
2970 LINS ALA THR GLY GLY ALA VAL THR LEU THR A A G C A A C G G G C G G A G C T G T G A C T T T A A C C 2980 3000	R THR GLU THR CACCGAAACAA 3030SER SER ASP ILE THR SER SER ASN GLY GINGTTCTGATATTACCTCAAGCAATGGTCAGGTTCTGATATTACCTCAAGCAATGGTCAG	P SER SER ILE TAGCAGTATCG 3090 ALA GLY SER ILE ASN ALA ALA ASN VAL THRCAGGAAGCATCAATGCCGCCAATGTGA.CA 3100 3110	
1970 TR GLY CGGGC 2980	THR A C A A 3030 SP ILE A T A T T 3040	TCG 3090 TR ILE CATC 3100	VAL 3 T G G 3150
2970 THR A C G G 2980	THR A A C A A 3030 ASP G A T A 3040	TATC (3090 SER 3100	. VAL T G T G 3150
ALA GCA	CGA CGA SER	CAGT CAGT GLY	THR TAC
IYS ALA	A A C C A C C G A A A C A A 3020 3020SER SER ASP ILEGTTCTGATATT 3040	S ASP SER SER ILE G G A T A G C A G T A T C G 3080 ALA GLY SER ILE C A G G A A G C A T C 3100	ASN THR THR GLY THR LEU THR THR VAL A A T A C C A C A G G C A C T T T A A C T A C T G T G G 3130 3150
2960	J TE A A C 3020	3080 3080	R LEU TTTA 3140
	T T T	A LYS	Y TH
	7 THE	A ALA	A GL
2950	.Y GL.) ; A G G 3010	10 THE CAC 3070	R THI CCAC 3130
	A G G	R LEG	N TH
	ALA THR GLY GLY THR LEUGCAACAGGAGGTACTT.	THR THR LEU THR ALA LYS ACAACTCTCACGGCCAA(3070	LEU ASN THR THR GLY THE TTAAATACCACAGGCAC 3130
	ALi	T刊 A C	E L

C	_	
	α	
٦	_	
7	_	
7	<u>۔</u> ح	

39/235		
THR LEU VAL ILE ASN ALA LYS ASP ALA GLY SER LYS ILE GLU ALA ALA SER GLY 3160 3170 3180 ACCCTGGTTATTAATGCAAAAGATGCTCAGT 3190 3200 3210 3220 3230 3240 GE	GLU VAL ASN VAL THR ASN ALA ASN GLY SER GAAGTAAATGTAACTAACGCAAATGGCTCCG 3270 3250GLY SER VAL ILE ALA THR THR SER SER ARGGCAGCGTAATCGCGACAACCTCAAGCAGA	VAL ASN IÍE THR GLY ASP LEU ILE THR ILE G T G A A C A T C A C A A T A A 3330 3310 3320ASN GLY LEU ASN ILE ILE SER LYS ASN GLYAT G G A T T A A A T A T C A T T C A A A A A C G G T 3340

PCT/CA99/00938

40/235

2	Υ
	α
7	
(5
Ĺ	工

							7	, , ,			
ILE	ATA	3600				ASN	AAT	3660			
GLU PHE ALA THR ARG PRO LEU SER ARG ILE	A A T T T G C A A C C A G A C C A T T A A G T C G A A T A					ASN SER ASP GLY ALA THR VAL CYS VAL ASN	A C A G T G A T G G C G C G A C G G T G T G C G T T A A T				
贸	A G T					CYS	TGC				
LEU	T T A	3590				VAL	GTG	3650			
PRO	C C A	(*)				强	A C G	(*)			
ARG	A G A					ALA	S C G				
黑	ACC	000	:	1 A	3630	GLY	\mathcal{C}	01			
ALA	GCA	3580	SER	TCA	363(ASP	GAT	3640			
出	T T T		出	GGGCGTGTTTCTCAA		段	A G T				
CID.	. A A	•	CYS	3 T G J		. ASIN	. A C	•			
:	:	•	ALA	000	3620	:	•	•	* *	GGTAG	3680
			ARG	AGG			-		ARG	9 9 J	
			GLY	0 9 9 1					GLY	G G G	
			CIN	GAA	3610				ASN	AAC	3670
			SE	TCI	36				ASP	GAT	36
			VAL ILE SER GLU GLY ARG ALA CYS PHE SER	GTGATTTCTGAAGGCA					ILE ALA ASP ASN GLY ARG ***	ATCGCTGATAACGGCC	
			VAL	G T G					ILE	ATC	

PCT/CA99/00938

<	\
C	2)
۲	_
(5
ī	L

Joyc hmw2A sequence

ASN	VAL	GLY
A A T	G T A	G G C
60	120	180
PRO ASP ASN VAL CGGATAATGTA 30 SER ILE GLU ASN PRO SER THR GLU ARG ASN TCCATTGAAAATCCTTCAACTGAACGCAAT TCCATTGAAAATCCTTCAACTGAACGCAAT TCCATTGAAAATCCTTCAACTGAACGCAAT	AGTATACAGGA AGTATACAGGA 30 90 THR GLY GLU ASN ILE ASN ASN PRO LYS VAL ACAGGGAAAATATAAACAACCCTAAGGTA 120	THR LEU THR SER CACTAACAAGC 150 SER ILE LEU GLU ASN ILE LEU LYS GLY TCAATCCTTGAGAAAAAGGC TCAATCCTTGAGAACATCCTGAAAAAGGC
THR GLU	PRO	LYS
ACTGAA(CCT	A A A
THR	ASN	LEU
ACT	A A C	CTG
50	110	170
SER T	ASN A A C	ILE A T C
PRO	ILE	ASN
C C T I	A T A	A A C
ASN VAL A A T G T A 30 ILE GLU ASN T T G A A A A T (THR GLY ACAGGA 90 GLY GLU ASN ILE ASN ASN GGAAAATATAAACAACC	TH LEU THR SER CACTAACAAGC 10 150 SER ILE LEU GLU TCAATCCTTGAGA
VAL G T A 30 GLU G A A	GLY GGA 90 GLU GAA	SER A G C 150 LEU GI
ASP PRO ASP ASN VAL GACCCGGATAATGTA 20 30 SER ILE GLU AS TCCATTGAAAA	LEU GLU TYR THR GLY C T A G A G T A T A C A G G A 80 90 THR GLY GLU AS A C A G G G A A A A	LYS THR LEU THR SER A A A C A C T A A C A G C 140 SER ILE LEU GI T C A A T C C T T G ?
ASP GATA SER TCCA	AGTATA AGTATA 30 THR ACAG	LEU CTA SER TCA
PRO	GLU	THR
C C G	G A G	A C A
20	80	140
ASP	LEU	LYS
G A C	CTA	A A A
LEU	ASP	LYS
T T A	G A C	A A A
LEU 3 T T G	GLU 1 G A A 70	1 SER GTCT 130
TRP	ASN	GIN
T G G	A A T	C A G
LYS GLU TRP LEU LEU	ASP SER ASN GLU ASP	ASN ASN GIN SER LYS
AAAGAGTGGTTGTTA	GATTCCAATGAAGAC	AATAATCAGTCTAAA
10	70	130
LYS	ASP	ASN
A A A	GAT	AAT

ILE ASN ILE GLY A T C A A C A T C G G A 230	VAL LYS ILE ASN GTTAAGATTAAT 290	TRP VAL ASP ILE TGGGTTGATATT 350	
SER TCTA	GLY GGTC	GLY GGAT	
SER AGCT	G G C G	LYS A A A Q	
THR ALA THR ASP ASN A C T G C C A C T G A T A A C 200 210 ILE TYR VAL ASN SER SER ILE ASN ILE GLY A T C T A C G T T A A T A G C T C T A T C A C A T C G G A A T C T A C G T T A A T A G C T C T A T C A C A T C G C A 240	ILE LEU SER GLY GLY ATTCTCTCAGGTGGA 270 GLY ARG ASN GLY GLY GLY VAL LYS ILE ASN GGCAGGAACGGCGGCGGTGTTAAGATTAAT GGCAGGAACGGCGGCGGTGTTAAGATTAAT	THR GLY GLY SER LEU A C G G C G G A G T T T A 320 THR ILE ASN SER LYS GLY TRP VAL ASP ILE A C C A T T A A T T C C A A A G G A T G G G T T G A T A T T 340 350	LEU GLY THR GLY PHE CTTGGTACGGGTTT 380
THR ACTG	ILE I	THR (LEU GLY
TAACI	TCACTT.	THR ACT 10	ASN ILE AATATTT 370
VAI T.	GE 3	LYS ASN ILE THR SER A A A A T A T T A C T T C C 310	HIS SER ASN ILE SER CACTCCAATATTTCA 370
PHE T T T G	SER	ASN A B A T	SER .
SER TCTT	ASP GACP	LYS A A A A	HIS C A C T

(ر	
_	$\overline{\mathbf{x}}$	
•	_	
	5	
Ĺ	L	

	44/235		
ALA G C T 420	GLN C A A 480	ASN A A T 540	ASP 3 A T 600
VAL 3 T G	ALA 3 C A	LEU	PHE T T C
SER T C C G	VAL 3 T C (SER I C I	LYS A A A A
GLY G G T T 410	ILE A T T 4	VAL G T G '	HIS C A T 7
ASN A A T	GIN C A A	ASN A A T	THR
LEU ASN ILE THR SER ASN GLY SER TTGAACATTACCTCTAATGGTTCC 400	ASP LYS ALA ARG A T A A G G C A C G T 10 450 SER ALA ALA ASP ALA GIN ILE VAL A G C G C A G A T G C T C A A A T T G T C 460 470	GLU ASN LYS GAAAACAAA 510 THR PHE ARG LEU ASN VAL SER LEU ASN ACCTTTAGCTTAACAATGTGTCTTTAAAT	ILE THR SER ATCACGTCA 570 ASN VAL GLY ASN GIN THR HIS LYS PHE ASP A A T T C G A T A A A T T C G A T 580 590 600
THR ACCI	 ASP G A T (LEU C T T 2	 ASN A A T
ILE A T A 400	ASP LYS ALA ARG G A T A A G G C A C G T 440 SER ALA ALA AS A G C G C G C A G P A G C G C G C A G P	GLY GLU ASN LYS G G G A A A A C A A A 500 THR PHE ARG LE A C C T T T A G G C 7	SER ILE THR SER TCCATCACGTCA 560 570 ASN VAL GLY AS A A T G T G G C A A 580
ASN A A C	ALA GCA ALA GCG(ASN A A C P PHE T T T P	THR ACG VAL GTG
	LYS A A G C SER A G C G	GLY GLU GGGAAA 00 THR ACCT	ILE ATCA ASN AATG
: : :	G 44		SER T C C 560
	LYS A A A	THR A C A	LEU CTA
	ASP G A C	LEU CTC	GLY G G T
	G G C A 430	A A A C 490	7 GLN T C A A 550
	LYS A A G	ILE ATA 4	GLY G G T 5
	PHE GLU LYS ALA ASP LYS TTTGAGAAGGCAGACAAA 430	GLY ILE ILE ASN LEU THR GGCATCATAAACCTCACA 490	GLY VAL GLY GLN GLY LEU GGAGTGGGTCAAGGTCTA 550
	PHE TTT	G G C	GLY G G A

PCT/CA99/00938

_	_)
<u> </u>	7	5
(Γ	5
È	Ī	-

LA C A 660	LSN A C 720	EU T T 780	
ALA G C 1 66(ASN A A C	LEU CT7	
THR	LYS A A A A	PRO C C & C	
THR ACAP	GIN C A A P	THR A C A	
PRO ALA CCTGCGA .650	VAL G T A (THR A C A 7	
PRO C T (ASN A A C (PRO	
ALA 3 C A C	LEU ASN VAL STTAACGTA(GLU 3 A A 750 ARG PHE GLY PRO THR THR 3 G C T T T G G C C C A A C A C A C A C A C A C A C	
 A C T G	TRP F G G 690 SER THR G T A C T C	 PHE T T C	: : :
THR A C A 630 GIN TH C A A A C	TRP T G G 690 SER TE	GLU G A A 750 ARG PF C G C T 7	GLY G G C 810
THR GLY ASN VAL THR ACTGGAAATGTAACA 620 630 ILE ASN GLN THR ALA PRO ALA THR THR ALA ATTAATCAAACTGCACCTGCGACAACGCA 640 660	TYR ASP SER TYR TRP TACGATTCATATTGG 680 680 ASN VAL SER THR LEU ASN VAL GIN LYS ASN ASN VAL SER THR LEU ASN VAL GIN LYS ASN 710	THR ACT ASN AT	VAL PHE PHE ASN GLYGTATTCTTTAACGGC800
ASN A A T C ILE A T T A	SER TYR T C A T A T ASN VAL A A C G T C	ARG THR CGCACT SER ASN AGTAAT	PHE T T T
G G A A 6620 A A	ASP G A T T 680 A	LYS A A G C 740 A	PHE T T C 800
THR ACTG 6	TYR T A C (ILE ATTP 7	VAL G T A
		PHE T T T	
ASN A A C A O	PHE T T T 1	THR ACC'	GLY GGA
GLY GLU ILE ASN ILE G G T G A A A T T A A C A T A 610	TYR TRP ASN PHE SER TATTGGAATTTTAGC 670	SER SER PHE THR PHE TCAAGCTTTACCTTT 730	ARG SER SER GLY GLY CGAAGCTCCGGAGGG 790
GLY GLU	TRP T G G A	SER AGC 1	SER AGCT
G G T	TYR T A T T	SER T C A A	ARG CGAA

Ų	L	j
C	7)
*	_	-
_	r	÷
	Ŀ	Į
ī	7	_

IG.19F

Ĺ	_)
C	Ī)
7	_	_
(r	;
Ì	Ī	_

	48 / 235		
ILE A T C 1260	ARG A G A 1320	ALA G C G 1380	GLU G A A 1440
ASN ILE THR AATATTACCA 550	SER T C C	SER VAL	SER GLU AGTGA
ILE A T T	PHE ALA SER TTGCTTCC	SER T C A	THR A C C
ASN A A A T 1250	PHE C T T T 1310	THR GLY 1 C C G G C T 1370	: GLU AGAG 1430
G G A	ASP GACT 13	THRACC	LYS GLY GLU THR AAAGGAGAGACCA 1430
LYS A A G	ARG C G C G	LEU CTAA	LYS A A A
SER SER ASP ILE LYS GLY ASN ILE THR ILE A G C A G T G A C A T C A A G G G A A A T A T T A C C A T C 1260	THR LEU LYS ALA A C A C T G A A G C C 280 HIS ASN SER PRO ARG ASP PHE ALA SER ARG C A T A A C A G C C C T C G C G A C T T T G C T T C C A G A G A C T T T G C T T C C A G A G A C T T T G C T T C C A G A G A C T T T G C T T C C A G A G A C T T T G C T T C C A G A G A C T T T G C T T C C A G A G A C T T T G C T T C C A G A G A C T T T G C T T C C A G A G A C A G A C T C G C G A C T T T G C T T C C A G A G A C A G A G	TGAATGTTGAA TGAATGTTGAA 10 1350 GLY ASN LEU THR LEU THR GLY SER VAL ALA GGAAATTTAACCCTAACCGGCTCAGTTGCG	SER ILE LEU ASN C C A T T C T T A A C 1410 ASP ALA THR PHE LYS GLY GLU THR SER GLU G A T G C T A C T T T T A A G G A G A G C C A G T G A G 1420 1420 1420
ASP GAC 12	THR LEU LYS ALA TACACTGAAAGCC 1280 1290 HIS ASN SER PF CATAACAGCC	C T T G A A T G T T G A A 1340 GLY ASN LEU TF G G A A T T T A A C 1360	J SER ILE LEU ASN T C C A T T C T T A A C 1400 1 ASP ALA THR PH G A T G C T A C T T T 1420
SER	THR LEU LYS CACTGAAA(80 HIS ASN CATAACA	ASN VALA A TGTTC	LEU TCT' ALA ALA
SER A G C A	LEU ACTO HIS	ASN G A A ' GLY G G P	ILE CAT' ASP GAT
	128	LEU THR LEU GLY ASN LEU TAACCCTTGGCAACTTGA 1330 1340 G	T T C C 1400
		· · · · · · · · · · · · · · · · · · ·	
	A A A T G	I GLY TGG	TAAT
	R THF : A A C 1270	IR LEU CCT 1330	(S GLN A B G G 1390
	S SB CTC 1	J 证明 AAC 1	E LYS
	LYS SER SER THR ASN VAI AAAGCTCAACAAATGT 1270	THR LEU THR LEU GLY AST 1CCTTAACCCTTGGCAA 1330	ASP ILE LYS GLY ASN LEG SATATTAAAGGTAACCT 1390
	LY A A	E O	AS 3. A

1620

园

AT

FIG. 19H

Ø G Ø ATA GGA 田田 K ⊟ CATTGAAGGC ASIN GLY A A 1490 드 H ASN A T Ø ď TTA AAC ں ASIN ASP 国 ACCGCCGA 亡 ⊱ IH STS \mathcal{C} GGTGG A A C... ACCAAT... A C C... A A 1600 GCTGAA. GLY Ø GGC ⊱ E E æ AT A AA AA TTC Н AAT. AGTAACAATAAT ⊱ GGA 드 出 HE ASIN ASS AA ⊢ AA Ø AAC CAAGGT AAAACC ASIN ASIN IH GLY ည် ဗ GLY GEN ASIN G CAA \mathcal{C} GIN 開 日日 S 関 Ø Ø ⊢ ⊱ ں ں 田田 ASN ASP ASIN Ø Ø Ø ACA . G AACA Ø ø ASN LYS ASN Ø Ø K G CAA G Ø 园 SIS T T \circ G AAC ည ပ C <u>ე</u> ASIN ASIN Ø

49/235

A A T

Н

CAA

A A

GIN
Ŗ
HE IE
ASIN
GLY
GLY
日

LYS	AAA	1680
GIN	TCCAAATTGGCGGCAATATCTCGCAAAAA	
Ŕ	TCG	
田	ATC	1670
ASN	AAT	` .
GLY	0 G C	
GLY	G G C	90
田	ATT	1660
GIN	CAA	
IE.	ATC	
:	:	:

:	:	
LYS	AAA	
ASP	G A T	
SER	TCTGATAAA	
SE	TTCT	
IE	⊱⊣	•
開	CTCACAA	
E E	CTC	
ASN	₽	
GLY	AAAGGCAA	
LYS	AAA	

			7
		開	ACA
		LYS ILE THR	ATA
		LYS	AAG
		HR LYS I	AAG
:	:	强	A C C
A A A.	1710.	TLE THR	TTI
GAT.		ASIN	ATCAATATTACCAAGAAGATAACA
T C T		ILE ASN	A T C i
TCT	1700	:	7
ATCTCACAATTTCTTCTGATAAA.	1690		

1	abla	
-	Ø	
	G	
1 1	Ø	
-	K	
,	\mathcal{O}	
1111	ACCAAGAA	
4	Ø	0
,	⊱	1720
1	₽	~
7	Ø	
-	⊱	
2	Ø	
4	Ø	
ר.	ATCAAT	
1	₽	
_	Ø	
:	:	:
:	:	:

A A A 1740

ATC.

:	SE SE SE SE SE SE SE SE SE SE SE SE SE S
1770	ALA
	PRO
	SER
1760	:
1750	

EE E

ALA

ASN

::)	1830
	1820
	1810
u	

GLY ້ ເ

GCA

PCT/CA99/00938

ASIN

H

	51/235		
ASN A A T 1920	ASN 1 A T 1980	ASN 1 A T 2040	
ALA Z	HIS A	SER A G T A	
ASP GLY ASN ALA BATGGTAATGCTA 1910	GLY 3 G T C	ASP 3 A T P	
GLY F G G T A 1910	ALA ASN GLY SCTAACGGT 1970	ALA ASP SCTGATO 2030	
ASP 3 A T (ALA G C T 7	ALA G C T (
SER AGTG	SER T C A G	SER A G T G	
ATTACAGCTAAAATGGCAACGATTTAACT 1870 1870 ILE GLY LYS ALA SER ASP GLY ASN ALA ASN ATTGGCAAGGCTAGTGGTAATGCTAAT 1920	THR PHE ASP LYS VAL LYS CTTTTGACAAGGTTAAA 1940 1950 ASP SER LYS ILEGATTCAAAATC7 1960	VAL THR LEU ASN SER LYS VAL GLU THR SER G T A A C A C T C T C T 2010 1990 ASN SER ASP SER ALA ASP ASP SER AST 2040 A A T A G T G A T A G T G A T A G T A A T A G	ASP ASN ASN THR GLY LEU THR ILE SER ALA GATAACAACTGGTTTAACCATTCCGCA 2050 2050
1870	LYS LYS VAL A A A A A G T G A 1930	LEU ASN TAAA 1990	ASN THR AACAC 2050
TTACAG	LYS A A A A	VAL THR LEU TAACACTA? 1990	ASN A A C A
ATTA	ALA G C T A	VAL G T A	ASP G A T A

_	<	•
C	T .)
T		-
(r	5
Ĺ	Ī	-

	52/235		
LYS ASP VAL THR VAL ASN ASP VAL THR A A G A T G T A A C A G T G A C G T C A C C A A G A T G T A A C A G T G A C G T C A C C A A G A T G T A A C A G T G A C G T C A C C	ILE SER ALA THR A T C T C T G C C A C A 2130 THR GLY ASN VAL THR THR LYS GLU SER THR A C A G G A A T G T A A C A A C C A A G A A G C A C A 2140 2150 2160	THR ILE ASN ALA ALA THR GLY SER VAL GLU CCATTAATGCGGCCACAGGTAGCGTGGAA 2190 2170 VAL THR ALA LYS THR GLY ASP ILE SER GLY GTAACTGCTAAAACGGCGATATTAGTGGC 2220	THR ILE SER GLY ASN THR VAL ASN VAL THR A C A A T T T C T G G T A A T G T T A C A 2250 2230 ALA THR ASP SER LEU THR THR GLN ALA SER G C A A C T G A T A G C T T A A C C C A A G C 2280
	HR ILE CAATAI	THR ILE ASN ALA ALA THR ACCATTAATGCGGCCACA 2170	THR ILE SER GLY ASN 1 ACAATTTCTGGTAATA 2230

ATGGTCAGACA ATGGTCAGACA 2310 THR LEU THR ALA LYS ASN GLY SER ILE ALA ACTCTTACAGCCAAGAATGGCAGTATCGCA 2320 2320 2320	SN VAL THR LEU A T G T G A C A T T A 10 2370 ASN THR GLY THR LEU THR THR VAL ALA A A T A C C A C G C A C C T T A A C T A C T G T A G C G 2380 2400	SER GLY THR AGTGCACT 2430 LEU ALA ILE ASN ALA LYS ASP ALA LYS LEU TTAGCTATTAACGCAAAGATGCTAAGTTA TTAGCTATTAACGCAAAGATGCTAAGTTA		
SER SER ILE THR SER SER ASN GLY GIN THR TCTAGCATTACCTCAAGTAATGGTCAGACA 2290 2310 THR LEU THR ALA ACTCTTACAGCCA 2320	GLY SER ILE ASP ALA ALA ASN VAL THR LEU GGAAGTATTGATGCCGCTAATGTGACATTA 2350 2370 ASN THR THR GLY ASN THR THR GLY 2380	GLY SER ASN ILE LYS ALA THR SER GLY THR GGTTCAAACATTAAGGCAACCAGTGGCACT 2410 2420 LEU ALA ILE ASN TTAGCTATTAACC	ASP GLY THR ALA SER GLY ASP ARG THR VALGATGGTACTGCATCAGGTGACGCACAGTA 2470	
SUBSTITUTE SHEET (RULE 26)				

GLY G G T 2520	ASN A A T 2580	ILE A T T 2640	ARG C G T 2700
VAL ASN ALA THR ASN ALA SER GLY SER GLY GTAAATGCAACTAACGCAAGTGGCTCTGGT 2500 2510	SER SER ASN VAL CAAGTAACGTG 2550 ASN ILE THR GLY ASP LEU SER THR ILE ASN AATATCACTGGAGATTTAAGCACAATAAAT AATATCACTGGAGATTTAAGCACAATAAAT AATATCACTGGAGATTTAAGCACAATAAAT	ASN GLY LYS AATGGTAAA 2610 ASN THR VAL VAL LEU LYS GLY ALA GLU ILE AACACCGTAGTGATAAAGGTGCTGAAATT 2620 2640	GLY VAL ALA GGTGTAGCA 2670 SER ALA ASN GLU VAL ILE GLU ALA LYS ARG AGTGCGAATGAGGTTATTGAAGCGAAGCGT
C C C	THR ACAA	ALA G C T	ALA G C G A
SER 1 A G T G 2510	SER A A G C P 2570	GLY A G G T C 2630	GLU C A A G 2690
ALA GCAP 25	LEU TTA	LYS A A A G 26	ILE ATT
ASN A A C G	ASP G A T T	LEU TTA?	VAL GTT?
THR ACTA OO	VAL 3 T G 2550 THR GLY . C T G G A 2560	LYS A A A 2610 VAL VAL T A G T G '	ALA 3 C A 2670 ASN GLU ASN GLU 2680
ALA G C A A 2500	VAL G T G 2550 THR GI A C T G (LYS A A A 2610 VAL V G T A G C	ALA G C A 2670 ASN GI A A T G A 2680
ASN A A T G	THR SER SER ASN VAL A C C T C A A G T A A C G T G 2540 2550 ASN ILE THR GI A A T A T C A C T G (2560)	SER LYS ASN GLY LYS TCGAAAATGGTAAA 2600 2610 ASN THR VAL V AACACCGTAG	GIN PRO GLY VAL ALA C A A C C A G G T G T A G C A 2660 SER ALA ASN GI A G T G C G A A T G A
VAL G T A A	SER A A G T A ASN A A T A	YS ASN A A A A T C 00 A A C A	GLY A G G T C SER A
: : :	SER T C A A 2540	LYS 3 A A A A 2 2600 A	PRO A C C A G 2660 A
	ALA G C A	ILE	ILE
	ALA A S C C C C C C C C C C C C C C C C C	N ILE ATATCA 2590	S TYR 1 A T A T 2650
	THR A ACTG 2530	ASN A A T 25	LYS SAAA 26
	SER VAL THR ALA ALA AGTGTGACTGCGGCA 2530	GLY LEU ASN ILE ILE GGATTAAATATCATT 2590	ASP VAL LYS TYR ILE GATGTGAAATATATT 2650
	SER A G T	GLY GGA	ASP G A I

FIG. 19N

	55/235	
GLY G G T 2760	GLU G A A 2820	SER A G T 2880
LEU CTT	ASN A A T (SER T C A A
LYS A A A	GIN C A A	PHETTCI
ALA A G C T 2750	THR 7 A C A 2810	CYS 3 T G T 2870
THR LEU ALA LYS LEU CATTAGCTAAACTTG 2750	ASN A A T	AIA G C G
THR , A C A	VAL 1 G T C	LYS A A G
ATTTATCTGAT ATTTATCTGAT 2730 GLU GLU ARG GLU THR LEU ALA LYS LEU GLY GAAGAAACATTAGCTAAACTTGGT 2760	VAL GLU PRO ASN TIGAGCCAAAT 80 2790 ASN THR ILE THR VAL ASN THR GLN ASN GLU AATACAATTACAGTCAATACACAAAATGAA AATACAATTACAGTCAATACACAAAATGAA AATACAATTACAGTCAATACACAAAATGAA	SER GIN VAL THR G T C A G T G A C A 40 ILE SER GLU ASP LYS ALA CYS PHE SER SER A T T T C T G A G A C A A G G C G T G T T T C T C A A G T A T T T C T G A A G A C A A G G C G T G T T T C T C A A G T A T T T C T G A A G A C A A G G C G T G T T T C T C A A G T A T T T C T G A A G A C A A G G C G T G T T T C T C A A G T 2860
ALA LEU GLU LYS VAL LYS ASP LEU SER ASP GCCCTTGAAAAGTAAAGATTTATCTGAT 2730 2710 2720 2730 CLU ARG GIGAAGAAGAGA	VAL SER ALA VAL ARG PHE VAL GLU PRO ASN G T A A G T G C T G T T T T G T T G A G C C A A A T 2790 2770 ASN THR ILE TH E T A C A A T T A C A A T T A C A A T T A C A A T T A C A A T T A C A A T A C A A T A C A A T A C A A T A C A A T A C A A T T A C A A C A C	PHE THR THR ARG PRO SER SER GIN VAL THR TTTACAACCAGACCGTCAAGTCAAGTGACA 2850 2850 ILE SER GLU AS ATTTCTGAAGA

ASIN GCAGTATGT GGTAATGGTGCA 2890 ALA ASN

G A T A G A C 2920

Ø

G

K1 hmw1A sequence

FIG.20A

57/235

ACT Ø ⊱ ACC ARG ASN TATAAAAA GGA AAC. GLY LYS $C \perp T$ AGA 国 ARG GCA ... TCGCCAAATCAAATTAAC AA ALA ASIN Ø C C C K PRO IE IEI E ₽ AACAATGAGTACGAC... CTAACAAACACAACA... TGCA TA ALA S AA 160 ASIN ASIN ARG A A A G ASN GACCCGGATAA ... TCTATT G A G 88 GEU ASP SE ... C T T EEI SE PRO ASP ASIN E ACA Ø ₽ ASN E IK AA ₽ G CCCA ⊱ 冒 SE GTT AAG R 图 T G A C <u>Г</u> AAAGAG CCATCC GAG

FIG.20E

VAL ASN ILE THR ALA THR LINS THR ILE THR GTTAATATCACTGCCACCAAACAATCACA 190	370 380 390
--	-------------

C	ر	
C	コ	
C	V	
,	.	
C	ני	
ī	_	
_	_	

	59/235		
PHE GLU ASP LIX SER ASN LEU ASN ILE THR THR LYS GLN GLY ASP ILE ALA 400 420 THE GLU ASP LYS PRO GLY LEU SER ASN LEU	440 450 450 THR ILE THR ALA LYS GLY THR ILE ALA VAL A C C A T T A C A G G G C C A T T G C G T G 480	ASN ASN LYS LYS GLY PHE ARG PHE ASP ASN A A C A A G A A G G C T T T A G G T T T G A T A A T 490 VAL THR LEU ASN GLY THR GLY GLY GLY LEU G T C A C T C T A A A T G G A G G G G C T C 540	SER PHE LYS TYR ILE GLU THR GLY ASN ARG TCTTTTAAATACATCGAAACCGGAAATAGA 550 ASP SER ASN PHE GLU THR HIS PHE ARG GLY GATAGCAATTCGAAACCCATTTAGAGGA GATAGCAATTTCGAAACCCATTTTAGAGGA

AAC 099 CAGT 780 ں r ⊱ TRP ပ ပ Ø 0 6 C C Ø ASIN Ø G A Ø GAG. CIN G G C G C T段 650 CAG GEN WAL G AAC G TTCCCT ARG ASIN PR0 A G TGGAAT... G G C... A T C... TATCG... Ø CGATTG B CAAGC ARG AGT Γ Ø G A T G ACC ... AGTGCC Ø 閨 G T A CGCTCCCAC C C C TTGATAGC T T A E ... G T A W 段 AAA ⊣ TTAAA G G G ARG HH IEO Ø GGA Ø G G T E 段 GLY ر 1 Ø E CTGGں ATCG CGT Ø ⊣ ASIN ASN C A Ø CA Ø Ø K AGA Ø ပ ပ ARG 图 G AGA S S S ۳ E ARG ARG ASN A A T A

ASN A A C 9

SER AGC

SER T C C.

ASN A A C

ALA SER SER ALA VAL ASN PHE ASN ILE LYS 920 840 940 940 950 870 940 940 940 940 940 940 940 9	ASN HIS THR LLEU PHE LYS GLY ASN ILE SER A A T C A C A T T A T T C A A G G G A T A T T T C A 910 VAL LEU GLY GLY GLY MET SER THR PHE ILE G T T T T A G G G G G G G A T G T C A A C T T T C A T T 940 950 960
---	---

S A 20	T A 30	A G 10	° ° °
PHE : LTTA 1020	PHE T TTA 1080	AL /	GLU (GLU (1200
ر ب		A G	L
ASI A A	ALA G C T	GIA C A	PHETT
SER GIN ASN CACAAAACT 1010	ALA	LEU ASN GLN VAL A TGAACCAAGTTG 1130	THR
SER T C A 1010	HIS C A C G 1070	LEU T T G 1130	ILE A T A 7 1190
TYR GLY VAL ILE ILE GLU SER GLN ASN PHE S ATGGCGTGATTATAGAGTCACAAAACTTTA 1020	SER LEU LYS PHE A G C T T A A A T T C A 1040 LYS SER GLU GLY SER THR HIS ALA ALA PHE 7 A A A G C G A A G G T T C G A C A C G C C G C T T T T A 1080	SER C A	LEU LYS LYS SER C T C A A A A A A G C C 1160 1170LEU ILE ALA ASN LYS ASN ILE THR PHE GLU (T T A T A G C C A A T A A A A C A T A A C C T T T G A £ G T T A T A G C C A A T A A A A C A T A A C C T T T G A £ G T 180 1180
ILE A T A G	SER LEU LYS PHE A G C T T A A A A T T C A 1040 1LYS SER GLU GLY SER A A A G C G A A G G T T C G A	ILE LEU ASN ALA A T T T A A A T G C C A 1100 11100THR GLY GLY ASN ILE C T G G G G C A A T A T A T	 5 LYS FAAA
ILE : 3 A T T A 1000	PHE I T C A 1050 GLY S.	LA C C A 11110 ASN IN A T A C	ier G C C 1170 ASN A A T A
G A 10(PHE TT (1000	ALA G C C 111. 1120 C A A 1120	SER A G C 117 ASN C A A 1180
VAL G T G A 10	LYS A A GLU G A	ASN A T GL) G G	LYS LYS A A A A A A A A ILE ALA A T A G C C
GLY G C	SER LEU LYS GCTTAAAA 1040LYS SER GLAAAGCGA	U A A SILY G G	S I A A A ILE T A
П.	LE S S S A A A	IE TT S (LY: A A A J J
ATGGC	SER A G C 1040 LYS A	ILE LEU ASN ALA A T T T T A A A T G C (1100 1THR GLY GLY AS C T G G G G C A A	LEU C T C A 1160 LEU T T
	SER SER LEU LYS PHE TCAAGCTTAAAATTCA 1040 1050 LYS SER GLU GLY SIAAAGCGAAGGTT	LEU ILE LEU ASN ALA T T A A T T T T A A A T G C C A 1100 1110 THR GLY GLY ASN II C T G G G G C A A T A T	ASN LEU LYS LYS SER A A T C T C A A A A A A A G C C 1160LEU ILE ALA ASN LT T A T A G C C A A T A A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C A A T A A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A A T A C C C A C C A C C C A C C C A C C C A C C C A C
	<mark>የ</mark> ካ	ASP G A T	
	GLY G G A 030	LYS ASN AAAAT C 1090	ASP
	ER ALA SER GLY GLY G T G C C T C A G G A G G (LA GLY ILE ASP SER CAGGTATTGATAGT 1150
	ALA G C C	HR ILE CAATAA	GLY G G T
	展 5 T 7	HR C A	C A

GCTT

TCA

AAA LYS

GAC

贸

63/235

A T A A A A... ASP GCAGCC ALA ALA L 园 開 Ā ATATC ASIN G 9 9 Γ_{N}

GE PR 8 ...IYS

ASN

Ø ₽ Ø ⊱ AA GGT AAA \mathcal{O} T Z AGAA A T ... A A C C A

 \mathcal{O} 1260

G T C A C C C T T C G T A... IEI IEI E ASN G G A Þ

Ø G GLY TYR ... SS:::

A A T (ASIN GGT ... G C G C G A A T T A T

AAAG... ⊱ ASIN Ø K ACT GTC M ₽ K Ø . წ

Ø

Ø

Ø

G

Ø

Ø

ARG A G

HH

1350... 開 园 ASIN · · · GLY

GCT JCC CTCACCGTTACCGGC AAT ... G C

<u>-</u>

A

SER

G T A... G Ø G A ACCGTT 国 ⊱ ASIN A A GAAAAA A T S

R S A C A 1440	SER ILE ALA LYS G CCATCGCTAAGG 1490	ER A C C G 1560	ASN I A A T A 1620
	T A 7	SE A T C	A A A
ASIN A A	ALA G C	LYS A A	LEU TT.
PRO C C A	ILE A T C	THRACT	ASP 3 A T
ALA ASN PRO ASN TYR S CTAATCCAAATTACA 1430 1440	SER T C C 7 1490	ILE THR THR LYS SER A TTACCACTAAATCCG 1550	GLY G G T (
ALA G C T	ILE A T T T	ILE A T T	LYS A A A A
LEU TTA	 70 R A SIN	 A 30 1 ASN G A A T	 ARG A G A
; PHE G T T T 1420	ASN 1470 S SER AS G T C A A A 1480	ASN A A T A 1530 LEU AS	CLY G G T A 1590 R ASN A T A A C A 1
LYS A A G	ASP A C A LYS A A G	GLU AGA SER AGT	LYS A G G A G G THR A C T
ALA GCT	PHE ASP TTGAC GLY LY GGCAA	ILE ATCG GLY GGC	ILE TAA ILE ATA
SER ALA LYS PHE LEU ALA ASN PRO ASN TYR :GTGCTAAGTTTTAGCTAATCCAAATTACA	GCCTATTGACAAC GCCTATTTGACAAC 1460 GIN GLY LYS SER ASN ILE SER ILE ALA LYS AAGGCAAGTCAAACATTTCCATCGCTAAGG AAGGCAAGTCAAACATTTCCATCGCTAAGG	LY GLY ALA ILE PHE LYS ASP ILE GLU ASN GAGGAGCTATTTTAAAGATATCGAGAATA 1510 1520 1530THR GLY SER LEU ASN ILE THR THR LYS SER A 1560 1560	HIS THR ILE ILE LYS GLY ATACTATTATAAGGGTA 1580 1580 ASN ILE THR ASN ARG LYS GLY ASP LEU ASN 1600 1610
	E C G	LYS A A A (色の
	ASN VAL SER ACGTATCC 1450	PHETTT	HIS C A T
	VAL C G T A 1450	ILE 1 A T T T T 1 1 1510	HIS C C A C C 1570
	ASN A A C	ALA 3 C T A 15	ASN A A C 1
	ER PHE ASN VAL SER CGCTTTAACGTATCCG	LY GLY GAGGAG	SP SER ASN HIS HIS TAR OF CONTAINS TO THE TRANSPORT OF TH
	日 い T	LY G A (SP A C T

FIG.201

LE THR ASN ASN GLY ASP ASN THR GLU ILE TCACGAATAATGGTGATAATACTGAAATCC 1650 1630 GIN ILE GLY GLY ASN ILE SER GLN LYS GLU GAAATTGGCGGCAATATCTCGCAAAAGAAG	LY ASN LEU THR ILE SER SER ASP LYS VAL G C A A T C T C T C T G A T A A G T C A 1690 1710 1.ASN ILE THR GLU ARG ILE THR ILE LYS ALA G A T A T T A C C G A G C G G A T A A C A A A G C A G A T A T T A C C G A G C G G A T A A C A A A G C A G	LY VAL ASN GLY ASP ASN SER ASP SER ASN G C G T T A A T G G G A T A A C T C T G A T T C A A A T G 1750 GLU ALA THR SER ALA ASN LEU THR ILE LYS T A G G C A A C A A G T G C T A A C C T A A A A A A A A 1780 1780	HR LYS GLU LEU LYS LEU THR ASN ASP LEU CCAAAGAGTTAAATTAACAAACGACCTAA 1810 1820 1830
LE THR ASN ASN GLY	LY ASN LEU THR ILE	LY VAL ASN GLY ASPGCGTTAATGGGGGA7	HR LYS GLU LEU LYS
TCACGAATAATGG	GCAATCTCACAAT		CCAAAGAGTTAAAA
1630	1690		1810

FIG.20J

	66/235		
ASN ILE SER GLY PHE ASN LYS ALA GLU ILE T A T A T T T C A G G T T T T A A T A A A G C A G A A A T T A 1840 1850 1860	LYS ASP ASN SER ASN LEU THR ILE 1 A A G A T A A C T A T T G 1870 1. GLY ASP ASN SER ASP ALA GLY ASN THR ASP A 1. G C G A T A A C A G T G A C G C T G G C A A T A C T G A C 1920 1920	LA LYS LYS VAL THR PHE SER ASN VAL LYS CTAAAAAAGTAACCTTTAGCAATGTTAAAG 1930 1ASP SER LYS ILE SER ALA SER ASP HIS ASN VASP ATTCAAAATCTCTGCTAGCGACCATAATGATTCAAAAATCTCTGCTAGCGACCATAATG	AL THR LEU ASN SER LYS VAL GLU THR SER TAACGCTAAACAGCGAAACATCTG 2010 1990 GLY ASP THR ASP SER THR GLU ASP GLY A GCGATACTGACAGCACTGAAGATGGCGGCA GCGATACTGACAGCACTGAAGATGGCGCA
	HR ALA I CAGCTA	LA LYS LYS CTAAAAAAG 19	AL THR I TAACGC
			. 7

SN ASN ASN THR GLY LEU THR ILE THR ALA ACAATAACACCGGCTTAACTATTACTGCAA 2050 2050Lys asn val thr val asn asn asn ile thr som the section and the section shows the se	ASN ILE THR ALA SER A A T A T C A C T G C G T C A G 2120 2130GLU ASN VAL THR LYS ALA GLY THR THR I A A A A T G T T A C C A C C A G G G C A C A A C C A 2150 2160	LE ASN ALA THR THR GLY SER VAL GLU VAL TTAATGCAACCACAGGTAGCGTAGAAGTAA 2190 2170Thr ALA LYS THR GLY ASP ILE LYS GLY GLY ICAGCCAAAACAGGTGATATTAAAGGTGGAACAGCCAAAACAGGTGATATTAAAGGTGGAA	LE GLU SER ASN SER GLY ASN VAL ASN ILE TTGAATCCAATTCGGGTAATGTAATTA 2230 2240 2250
---	--	---	---

FIG.20L

	68/235		
THR ALA SER GLY ASP THR LEU ASN VAL SER ACAGCGAGCGACACGCTTAATGTAAGTA	VAL THR VAL ALA G T G A C A G C A G 2310ALA ALA SER GLY ALA VAL THR THR THR LYS G C A G C C T C A G G T G C C G T A A C C A C A A A A G C A G C C T C A G G T G C C G T A A C A A C A A A A G C A G C C T C A G G T G C C G T A A C A A A A A G	ASN ALA THR THR GLY ASN A A T G C A A C T G G T A A T G 2360 2370 ALA ASN ILE THR THR LYS THR GLY GLU ILE A C A A A T A T T A C A A C C A A A C A G T G A A T T A 2380 2390 2400	LYS SER ALA SER GLY ASN A A A T C A G C T T C C G G T A A T G 2420 2430 VAL ASN ILE THR ALA SER GLY ASN THR LEU A T A A A T A T T A C A G C G A C C G C A T A C A C T T A 2460
	SN ILE THR GLY GLN ASN ACATCACAGGTCAAAT 2290	LY SER THR ILE ASN ALA GATCAACTATTAATGCA 2350	SN GLY GLU VAL LYS SER A T G G C G A A G T T A A A T C A 2410

-1G.20M

4	_	_	
)	
C	`	L	
		3	
(ı)	
-	-	-	
L	1		

70/235		
SER SER LYS ILE AGCTCTAAGATTA 2720ASN GLY THR LYS SER VAL THR SER SER GACGGAACTAAGAGTGTAACCACCTCAAGCCACGGAACTAAGAGTGTAACCACCTCAAGCCACGGAACTAAGAGTGTAACCACCTCAAGCC	GLY ASP ILE SER GLY THR ILE SER AGGTGATATTAGTGCACACAATTTCTG 2770 2770GLY ASN THR VAL SER VAL SER ALA THR GLY SGTAATACGGTAAGCGTTAGTGCGACCGGTAGTAATACGGTAAGCGTTAGTGCGACCGGTA	ER LEU THR THR GLN ALA GLY SER LYS ILLE GCTTGACCACTCAAGAAAATTG 2830 2840 2.830GLU ALA LYS THR GLY GLU ALA ASN VAL THR SAGCAAAACAGGTGAGGCTAATGTAACAA 2860 2870 2880
YS LEU AATTA <i>i</i>	IN SER AATCA(ER LEU G C T T G A
	LEU THR CLN THR SER SER LYS ILE TAACCACACAACAAGATTA 2730 2710 ASN GLY THR LYS SER VAL THR SER SER GASN GLY THR LYS AGAGTGTAACCACCTCAAGCC 2760	THE THE GIN THE SER LYS ILE

FIG.200

ASN VAL THR ALA ASN T AATGTTACAGCAAATA 2930 2940	CGGAGCTGTGACTT 2990 3000	THR SER SER ASN G TACCTCAAGCAATG 3050	
ER ALA THR GLY THR ILE GLY GLY THR ILE G C G C A C A G G T A C A A T C G G T A C A A T C T 2890 2900 SER GLY ASN THR VAL ASN VAL THR ALA ASN TA A T A C A G C A A A T A C A G C A A A T A 2920 2940	HR ASP ASN LEU THR ILE LYS ASP GLY ALA C T G A T A A T T T A A G A T G G C G C A A 2950 2950ARG ILE LYS ALA THR GLY GLY ALA VAL THR I GLY GLY ALA VAL THR I SP300 3000	EU THR ALA THR GLY GLY THR LEU THR THR TAACCGCAACAGGAGGTACTTTAACCACG 3030 3010GLU THR SER SER ASP ILE THR SER ASN (AAACAAGTTCTGATATTACCTCAAGCAATG	LY GIN THR LEU THR ALA LYS ASP SER GTCAGACAACTCTCACGGCCAAGGATAGCA 3070 3080 3080

PCT/CA99/00938

T G

A A

ာ ပ G

AAT

C

Ø

æ Ø

K

G

Ø

ا ا Z,

æ

ACT

 \mathcal{O}

CA選

Ø \mathcal{O}

Ø \vdash

CAØ

AL T G

园

GLY G G

ASIN Ø K K

LEU ₽

日日

SER

GLY . 9

ALA

ILE

...SER .: G

1///	7	2/	¹ 2	3	5
------	---	----	----------------	---	---

	7	72/2	235												
	ഗ	Æ	0				H	()	0				S	~	
	ALA	J J 5	3180				ASP	GAC	3240				SER	TCAA	3300
	ALA	GCA					GLY	GGT					訊	ACC	
	GLU	GAG					SER	T C A					開	A C A	
	IE	ATC	3170				ALA	G C A	3230				ALA	3 C G	3290
	LYS	AAA					ALA	0 C G					ILE	A T C (
:	段	T C A		•	:	:	GLY	J 9 9		:	•	:	VAL	GTA 2	
3150	GLY	GGT	3160	SP.	A T G.	3210	ASP	GAC	3220		ACG.	3270	SER	AGC (3280
	ALA	GCA	m	YS A	A A G		EEI	T T G	C	LA A	CAA		GLY	2 G C	m
	VAL	GTG		I WI	CAA		GIN	C A G		SNA	A C G		SER	I C C (
3140	THR VAL ALA GLY SER LYS ILE GLU ALA ALA S	CTGTGGCAGGTTCAAAAATCGAGGCAGCCA	:	ASN ALA LYS ASP	AATGCAAAAGATG	3200	ALA GIN LEU ASP GLY ALA ALA SER GLY ASP H	CTCAGTTGGACGGCGCGATCAGGTGACC	:	THR ASN ALA ASN	ACCAACGCAAACG	3260	GLY SER GLY SER VAL ILE ALA THR SER	G C T C C G C A G C G T A A T C G C G A C A A C C T C A A	:
				ILE						ALA					
				VAL	G T T					ASIN	AAT				
3130				GLY THR LEU VAL ILE	C T G	3190				THR VAL VAL ASIN ALA	GTA	3250			
ω,				採	ACC	m				VAL	GTA	ς.			
				GLY	GGCACCCTGGTTATT					出	ACAGTAGTAAATGCA				

G T

IS A C

	73/235		
ER ARG VAL ASN IIE THR GLY ASP LEU IIE G C A G A G T T T A A T C A 3330 3310 THR IIE ASN GLY LEU ASN IIE IIE SER LYS A C A A T A A A T G G A T T A A A T C A T T C A A A A 3360	GLY LYS ASN THR VAL LEU LEU LYS GLY GTAAAACACCGTGCTGTTAAAAGGTG 3390 3370VAL GLU ILE ASP VAL LYS TYR ILE GLN PRO GTGAAATTGATGTGAAATACATTCAACCGG	LY ILE ALA SER VAL ASN GLU VAL ILE GLU GCATAGCGAGCGTAAATGAAG 3430 3.430ALA LYS ARG ALA LEU GLU LYS VAL LYS ASP LCGAAACGCCCTTGAGAAGTAAAGATT 3.480	EU SER ASP GLU GLU ARG GLU THR LEU ALA TATCTGACGAAGAAACATTAGCTA 3490 3500 3510
ER G C A	SN A C G	LY G C A	EU T A

A G

G A G (3680

3670 T

₽

Ø Ø

	•	,, 200		
LYS LEU GLY VAL SER ALA VAL ARG PHE ALA G AACTTGGCGTGAGCGTACGTTTGCTG 3520 3530 3540	LU PRO ASN ALA ILLE THR ILLE ASN THR A G C C A A A T A A T C G A T T A A T A C A C 3550 3550GIN ASN GLU PHE THR THR ARG PRO LEU SER G	A A A A T G A G T T T A C A A C C A G A C C A T T A A G T C 3580 3580 3600	IN VAL THR ILE SER GLU GLY LYS VAL CYS A A G T G A C A A T T T C T G A A G G T A T G T T 3610 3620 3630 T.PHE LEU ILE GLY ASN GLY ALA THR ILE CYS T	3640 3650 3650 3660

FIG.21A

K21 hrm/14 sequence

A T A... ₽ G B ₽ ASP GACCCGGA 82 ASP Ø 园 L G 国 E. - 9 - 9 - 1 G G

Ø

A A

SEC GLY ASIN ⊣ A T ... A A T

75/235

L C C T

Ø . ق

ATT

AAT.

AGT

Ø

ტ ტ

ں

AA

ں

G

ILE

ACACCC... CAGCCAGGTAGAGGCGAT ASP GLY ARG GLY PRO GIN

Ø

Ø

ø

国

国 0 0 0 GAA GCA AAGGTTTCA 段 ... A A C ASIN

T C

ACA

Ø

段 EEG 開 ALA ASIN

TCCACCGCTTTA... TTA CA Ø ည ည E G Ø Ø AC Ø ₽ Ø

ASIN

 $C \subset C$ ILE ASIN WAL GEU GLY GLN

 $T \subset T$ ATT AAC GTO GAG ... CAAAAGGGTATT 160

76/235

ASN ALA ASP VAL C A A C G C G G A T G T T 200 210 ASP VAL LYS ASN GLY THR LEU VAL LEU HIS G A T G T T A A A A C G G A A C A T T A G T A T T A C A T 220 230 240 LYS ILE ASN GLY T A A A A T T A A C G T 260 270 ASN ILE THR SER THR GLN ASN GLY ASN LEU A A T T A C C T C A C C A A A A T G G T A A T T T A A A T T A C C T C A C C A A A A T G G T A T T T A A A T T A C C T C A C C C A A A A T G G T A T T T A A A T T A C C T C A C C C A A A A T G G T A T T T A A A T T A C C T C A C C C A A A A T G G T A T T T A A A T T A C C T C A C C C A A C C T G G T A T T T A A A T T T A C C T C A C C A A A A T G G T A T T T A A A T T T A C C T C A C C A A A A T G G T A T T T A A A T T T A C C T C A C C A A A A T G G T A T T T A A A T T T A C C T C A C C A A A A T G G T A T T T A A B T T T A C C T C A C C A C A A A A T G G T T T T A A B T T T A C C T C A C C A A A A T G G T A T T T A A B T T T A C C T C A C C A A A A T G G T A T T T A A B T T T A C C T C A C C A A A A T G G T T T T A A B T A T T A C C T C A C C A A A A T G G T A T T T T A A B T A T T A C C T C A A C C T T G G T A T G G T T T T T T T T T T T
--

390... : G A T A A T A 380 TCC ILE 1112.
TATTACTT 370 AAT LEU

FIG.210

THR . C C 420	SER CT 480	LYS , A G 540	ASIN 1 A C
T A	S T	L	A
ILE A T T	VAL G T A 1	HIS CATA	HIS C A C
PHE GLU LYS GLY ASP ASN LEU THR ILE THR TTTGAAAAGGTGATAATCTAACCATTACC 400 420	CTCTAATCAAGAG 440	GGGTTTAACTTTT GGGTTTTAACTTTT 500 510 THR ALA ASN LYS GLY ASN HIS THR HIS LYS ACTGCAAATAAAGGTAATCATACCCATAAG 520 530	CATTTCCGGAAAG 5560 570 VAL VAL ILE ASN GIN THR THR PRO HIS ASN 5
LEU C T A A 410	SER A G T A 470	HIS CATA	THR THR CCACAC
ASN AAT(PHE TTTA	ASN A A T C	THRACC
ASP GATA	A G A T	GGT	GIN CAA
PHE GLU LYS GLY TTGAAAAAGGT	LA GIN GLY ASN ILE ILE SER ASN GIN GLU CCCAAGGAATATAATCTCTAATCAAGAG 450 430 ASN LYS GIN LEU ASN ACAACTTA	EU ASN GLY MET GLY ALA GLY LEU THR PHE TAAATGGGATGCGGGTTTAACTTTT 490 510 THR ALA ASN LYS ACTGCAAATAAAG	HE ASP GLY THR LEU ASN ILE SER GLY LYS TTGATGGCACGCTTAACATTTCCGGAAAG 550 570 AN VAL ILE ASN GLN GTAGTAATTAATCAAP
LYS A A A 4	JIN A A G GIN C A A	THR CTT ASN AAT	ILY GAA ILE ATT
GLU GAA	LYS LYS A A A	GLY LEU THR GGTTTAACTT 000 THR ALA ASN ACTGCAAAT	SER GLY CCGGA VAL ILI GTAAT
PHE I T T	ER P	CY IST TO THE THE TABLE	ILE S TTT 60 VAL GTA
	C T C T 440 ASI	G G G G G S S S S S S S S S S S S S S S	C A T 560 V G
	ILE A T C T 4	AIA G C G C	ASN A A C
	ILE A T A	GLY 3 G T	LEU
	ASN ILE AATAT	META TGO	THR ACG(
	NIA GIN GLY P CCCAAGGAA 430	GLY N 3 G G A 490	ASP GLY 1 1 A T G G C A 550
	GIN 'AAG	ASN A A T G	ASP A T G
	EA C C C	EU A A	T T G

SUBSTITUTE SHEET (RULE 26)

78/235

720 GGT 099 Е Ø ⊣ Ø AAA 园 GTAGCT (710) E-G 用用 ALA MET ပ ď Ø C T TGAA LEG VAL GE 650 ACT $T \subset G$ A T 段 ASIN Ø AAT ⊣ AGCAACCGC ASN AC $T \subset T...$ Ø A A G... GGCAAG LYS T G T 760 贸 AA A A GTA GAA ASP G A T G G Ø TAC GTCGAT Ø GTT TTTGCGGGG ALA <u>ပ</u> ဗ TAC CCTTTA L TYR AGT AAT 069 段 CA ₽ ALA ည ဗ G Ø ⊱ ₽ H E ASIN ASN A G A A Ø Ø <u>ტ</u> A A AAGA ⊣ GEN TGA ₽ ڻ G Ø 88 段 ں $^{\circ}$ ⊱ C G Ç Ε¥ ⊱ <u>ი</u> G Ø AGC ⊱ ILE Ø E Ø Z Ø

FIG.21E

	79/23	35	
LYS LEU LYS SER ASN ASP ASN THR SER ASN A A G T T A A A A T C A A T G A T A C A A G C A A C A 820 840	별 e	CE	ARG GLY ASN ASN ALA CGCGTAATAATGCT1

PCT/CA99/00938

FIG.21F

GLU ILE LYS LYS ASP LEU ILE ILE ASN ALA GAAATCAAAAGATTTAATTATAATGCA 1040 1050 THR GLY SER ASN PHE ASN LEU LYS GLN THR THR GLY SER ASN PHE ASN LEU LYS GLN THR ACTGGCTCGAATTTAATCTTAAGCAAACG 1080	LYS ASP LYS PHE ASP ASN SER TYR GLU LYS A A G A T A A A T T T G A C A A T A G T T A C G A A A A 1090 ASN ALA ILE PHE SER THR HIS ASN LEU THR A A C G C C A T T T T C T C A T A A C C T A A C C A A C G C C A T T T T C T C A T A A C C T A A C C 1140	ILE LEU GLY GLY ASN VAL THR LEU GLY GLY A T T C T T G C C G C A A T G T T A C T C T A G G T G G 1150 1150 GLU ASN SER SER ASN ILE LYS GLY ASN GLU ASN SER GT A G T A A T T A A A G G A A A T T C A A G T A G T A T T A A A G G A A A T T C A A G T A G T A T T A A A G G A A A T T C A A G T A G T A T T A A A G G A A A T T C A A G T A G T A T T A A A G G A A A T T C A A G T A G T A T T T A A A G G A A A T T C A A G T A G T A T T T A A A G G A A A T T C A A G T A G T A T T T A A A G G A A A T T C A A G T A G T A T T T A A A G G A A A T T C A A G T A G T A T T T A A A G G A A A T T C A A G T A G T A T T T A A A G G A A A T T C A A G T A G T A T T T A A A G G A A A T T C A A G T A G T A T T T A A A G G A A A A	ILE ASN ILE ASN SER LYS ALA ASN VAL THR A T C A A C A T C C A A G G C A A A T G T T A C A
---	--	--	---

SUBSTITUTE SHEET (RULE 26)

(5
7	N
<u> </u>	<u>5</u>
Ĺ	匸

LEU C T T 1260	1LE A T A 1320	GLN C A A 1380	ASP G A C 1440
HIS C A C	GLY ASN LEU ASN ILE GAAATTTAAACAT 1310	PHE GLN TTTCA 7	ASN GLY THR ALA ASP ACGCACCGCCGA 1430
SER AGT	LEU T T A	ALA LYS S C T A A A T	THR
THR C A C G 1250	ASN A A A T 1310	ER ALA 3 T G C T 1370	N GLY CGGC 1430
O D D	GLY G G A	SER AGTG	ASN AAC
ALA G C C	9999 	GLU GAA	. ASN A A C
LEU GLN ALA HIS ALA GLY THR SER HIS LEU T T A C A A G C T C A T G C C G G C A C G A G T C A C C T T 1240 1260	THR LEU THR LEU GLY ACCTAACCCTTGGC 1280 1290 ASN VAL SER VAL GLY GLY ASN LEU ASN ILE ASN VAL SER VAL GLY GLY ASN LEU ASN ILE AATGTATCTGTTGGGGGAAATTTAAACATA 1320	HIS ILE ASP GLY ASN CATATTGACGCAAT 1340 1350 LEU SER ILE ALA GLU SER ALA LYS PHE GIN CTTTCTATTGCAGAAGTGCTAAATTTCAA 1380	GLY LYS THR ASN ASN LEU ASN ILE THR GGAAAACCAATAACTAATTAC 1390 GLY THR PHE THR ASN ASN GLY THR ALA ASP GGCACCTTTACCAACACGGCACGGCGAC 1440
ALA AGCT 12	LEU GLN 2 T T G G 129 SER V 1300	GLY AST 3 G C A A 135 ILE P 1360	ILE THE THE 147 PHE T T T A C 1420
GIN A C A A	THR ACC(VAL	ASP 3 A C (SER I T C T	ASN A A T 7 THR
	LEU C C T A . 1280 ASN	S ILE TATT 1340 LEU	C C T A 1/1400 GLY G G C
: : :	THR ACC 1 1	HIS C A T 1	ASN A A C 1 1
	ARG A G A		ASN A A C
	S GLU A A G A A A 1270	SER ASN CAAATO 1330	ASN A A T A 90
	LYS A A A A A A A A A A A A A A A A A A A	SER T C A 13	THR 7 ACCA 1390
	ASP LYS LYS GLU ARG G A T A A A A A G A A A G A 1270	ILE GLY SER ASN ALA ATTGGCTCAAATGCA 1330	GLY LYS THR GAAAAACCA 1390
	ASP G A 1	ILE AT 1	GLY GG&

ILE ASN ILE LYS GLM GLY VAL VAL LYS LEU A T T A A T A T A A A C A A G A G T G G T A A A C T C 1460 GLM GLY ASP ILE THR ASN ASN GLY ASN LEU C A A G G T G A T A T T A C C A A T A A T T A T T A T T A T T A T T A T T A T T A T T A T T T A T T T A T T T A T T A T T A T T A T T T A T T T A T T T A T T T A T T T A T T T A T T T A T T T A T	ASN ILE THR THR ASN ALA SER VAL ASN GIN A A T A T C A C T C A G T C A A T C A A 1510 LYS THR ILE ILE ASN GLY ASN ILE THR ASN A A A A C C A T T A T T A A C T A A C 1560	LYS LYS GLY ASP LEU ASN ILE LYS ASP ILE A A A A A A G C C A C T T A A C A T C A A G A T A T T 1570 LYS ALA ASN ALA GLU ILE GLN ILE GLY GLY A A A G C C A A C G C C G A A T T G G C G C 1620	ASN ILE SER GLN LYS GLU GLY ASN LEU THR A A T A T C T C G C A A A A A A G G T A A T C T C A C G
--	--	---	--

LYS A A A 1680	ASN A A T 1740	GLY G G T 1800	SER A G C 1860
THR ACCA	SER A G T	SER T C A (ALA 3 C T
ILE A T C A	ALA 3 C A A	ILE \TTT	LYS A A G (
ASN 1 A A T A 1670	VAL C G T A (ASN A A C A 790	GLY 1 G G C P 1850
ILE TTA	GLY VAL GCGTA(LEU ASN CTAAA 1790	ILE 1
SYS A A A	SER C	ASN	ILE .
ILE SER SER ASP LYS ILE ASN ILE THR LYS A T T C T C T C A A A A T T A A T A T C A C A	ASP THR ASP GIN AGATACTGATCAA 1700 1710 GLY ASN SER ASP SER GLY VAL ALA SER ASN GGGAATTCTGATTCAGGCGTAGCAAGTAAT 1740		A :
SER P I C T G 1660	ALA ASP THR ASP GIN GCAGATACTGATCAA 1700 1710 GLY ASN SER ASP GGGAATTCTGAT	LYS THR LYS GLU LEU A A A C C A A A G A G T T A 1760 1770 THR LEU THR ASP A C A T T A A C A G A C I	ILE THR ALA LYS ASP A T T A C A G C T A A A G A T 1820 1830 ASN SER ASP LEU A A C A G T G A T T T A A
SER ICIT	ASP THR AS AGATACTGA 1700 GLY ASN GGGAATT	YS GAAGAAGA	LA LY CIAN
ILE \ T T T	ASP TISTA (100) GLY G G G G	R LY CAA	HRAD CAGO
 	ALA ASP ; C A G A T 1700 GLY G G (A A C C 1760 1760 1760 1760 1760 1760 1760 1760	ILE THR T T A C P 1820 AS A A
	LYS A A G	ILE A T T	GLU 3 A A
	ILE A T T A O	THR ACCA	ALA GLU SCAGA, O
	ARG ILE GLU ILE LYS CGGATAGAAATTAAG 1690	ALA ASN LEU THR ILE GCTAATCTAACCATT 1750	PHE ASN LYS ALA GLU TTTAATAAAGCAGAA 1810
	ILE A T A G	ALA ASN LEU CTAATCTA 175	ASN A A T 7
	ARG C G G A	AIA G C T	PHE T T T.

FIG.21J

TAAACAAATAACC 1880 PHE ASP LYS VAL LYS ASP SER LYS ILE SER TTTGACAAGGTTAAAGATTCAAAAATCTCA 1920	A A C A C T A A A T A G C 1940 LYS VAL GLU THR SER ASN SER ASP GLY SER A A A G T G G A A C G T C T A A T A G C G A T G G T A G C A A A G T G A A C G T C T A A T A G C G A T G G T A G C A B A G T G A A C G T C T A A T A G C G A T G G T A G C	TGACAACATATC 2000 2010 GLY LEU THR ILE SER ALA LYS ASP VAL THR GGCTTAACTATTTCCGCAAAAGATGTAACG	THR A C A
SER ASP ASN SER ASN ALA LYS GIN ILE THR A G T G A C A G T A A T G C T A A C A A T A A C C 1890 1870 PHE ASP LYS VAL T T T G A C A A G G T T A	ALA GLY ASN HIS ASN VAL THR LEU ASN SER G C T G G C A A T C A C A A T G T A A C A C T A A A T A G C 1930 LYS VAL GLU THR SER A A A G T G G A A A C G T C T A A A B C G T C T A A A B C G T C T A A A B C G T C T A A A B C G T C T A A A B C G T C T A A A B C G A A A C G T C T A A A B C G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C G T C T A A A B C T G A A A C T C T A A A B C T G A A A C T C T A A A B C T G A A A C T C T A A A B C T G A A A C T C T A A A B C T G A A A C T C T A A A B C T G A A A C T C T A A A B C T C T A A A C T C T A A A B C T C T A A A A B C T C T A A A B C T C T A A A A B C T C T A A A A B C T C T A A A A B C T C T A A A A B C T C T A A A A B C T C T A A A A B C T C T A A A A A A A A A A A A A A A A	THR GLY ASN GLY SER ASP ACCGGAAACGGTAGCGA 1990	VAL ASN SER ASN ILE THR SER HIS LYS THR GTAAATAGTAATATCACCTCTCACAAAACA 2070

ILE A T C 2100	LYS A A A 2160	ASP G A C 2220	THR. A.C.T 2280
GLY G G T	ALA G C T	THR A C C	LEU
GLY 3 G A	THR ACTG	THR ACCA	THR CCT
GLU A G A A (2090	VAL A G T A A 2150	ALA 1 G C A 7 2210	ALA THR LEU FGCAACCTTA 2270
SER GLU GLY GLY CAGAAGGAGGT 2090	GLU 3 A A G 21	THR ALA CAGCAA 2210	THR C T G
ALA 3 C A T	VAL 3 T G G	VAL	GLY '
) C T (()	(')	
ILE S TCT 2080	CACAACCATTAATGCG 2120 2130 THR THR GLY SER ACCACAGGTAGC	CGGTACGATTTCCGGT CGGTACGATTTCCGGT 2180 2180 LYS THR VAL SER A A G A C A G T A A G T C	GLY ALA LYS ILE T G G C G C A A A A T T 2240 ASN ALA THR GLU A A T G C G A C A G A A G
ASN ASIA A	THR THR ILE ASN AI CAACCATTAATGC 2120 21 THR THR GLY ACCACAGGTA	THR ILE SER GI TACGATTTCCGG 2180 21. LYS THR VAL AAGACAGTAA	A LYS AAAA LA THI CGAC
A A A	ILL A T C A	ILI A T S T G A	ALA GCA I ALA TGC
G T A A	A A C C 2120 THE	THR TACGA 2180 LYS AAG	GLY GLY G T G G C G 2240 ASN A A T
: : :	THR A C A 2	GLY 3 G T A 21 	GLY 3 G T 2 2
	:>-I	℃ .	LYS A A A (
	ALA GLY 3 C A G G (ILE SER ATTAGO	VAL LYS
	LYS A A A A A G 2110	ASP I 3 A T A 2170	THR 1
	THR THR CTACT?	GLY GLY G	LEU THR FTAACT 223
	THR THR LYS ALA GLY ACTACTAAAGCAGG 2110	THR GLY ASP ILE SEIACAGGCGATATTAG 2170	SER LEU THR VAL LYS GLY GLY ALA LYS ILE AGTTTAACTGTTAAAGGTGGCGCAAAATT 2230 2240 2250 ASN ALA THR GLU AATGCGACAGAA 2260

A A C C A C C G A G G C C 2300 2300 2310 ASN SER ALA ILE SER GLY ALA ASN GLY VAL A SU C G G G A T T A G C G G G G C T A A C G G T G T A 2320 2340	GLY ASP ILE SER AGGCGATATTAGC 2360 2360GLY THR ILE SER GLY LYS THR VAL SER VALGGTACGATTTCCGGTAAGACAGTAAGTGTT 2380 2400	TTTAACTGTTGGA 2420 2420 GLY ASP ALA LYS ILE ASN ALA THR GLU GLY GGTGACGCAAAATTAATGCGACAGAAGGA GGTGACGCAAAATTAATGCGACAGAAGGA	C A C T
ALA SER SER GLY LYS LEUGCATCATCGGGCAAATT 2290	THR ALA SER GEN SER GLY ASP ILE SER A C T G C C T C A A G T C A A T C A G G C G A T A T T A G C 2350 GLY THR ILE SER G G T A C G A T T T C C 2380	THR ALA SER SER GLY SER ACAGCAAGCTCTGGCAG 2410	ALA ALA THR LEU THR ALA THR LYS GLY THR GCTGCGACTTTAACTGCAACAAAGGCACT 2470 2480

FIG.211N

LEU THR THR VAL LYS GLY SER ASN ILE ASP TTAACTACCGTGAAGGGTTCAAACATGAC 2520	ALA ASN GLU GLY THR LEU VAL ILE ASN ALA GCAAACGAAGGCACCTTAGTTATTAACGA 2550 GIN ASP ALA THR LEU ASN GLY ASP ALA SER CAAGACGCACACACTAAATGGTGATGCATCA	GLY ASP ARG THR GLU VAL ASN ALA VAL ASN GGCGACCGTACAGAGTGAATGCAGTCAAC 2600 2600 ALA SER GLY ASN VAL THR ALA LYS GCAAGCGCTCTGGTAACTGCGAAA	THR SER SER VAL ASN ILE THR GLY ASP ACCTCAAGCAGTGTGAATATCACTGGAGAT 2650 2670 LEU SER THR ILE ASN GLY LEU ASN ILE ILE LEU SER THR 16 GATTAAATCATT 2680 2700
--	--	---	--

WO 00/20609

SER LYS ASN GLY LYS ASN THR VAL VAL LEU T C G A A A A T G G T A A A C A C C G T A G T G T T A 2730 2710 LYS GLY ALA GLU ILE ASP VAL LYS TYR ILE A A A G G T G C T G A A T T G A T A T A T T 2760	GIN PRO GLY VAL ALA SER ALA ASN GLU VAL CAACCAGGTGTAGCAAGTGCGAATGAGGTT 2770 ILE GLU ALA LYS ARG ALA LEU GLU LYS VAL ATTGAAGCGTGCCTTGAAAAGTA 2820	LYS ASP LEU SER ASP GLU GLU ARG GLU THR A A A G A T T T A T C T G A T G A A G A G A A C A 2830 LEU ALA LYS LEU GLY VAL SER ALA VAL ARG T A G C T A A A C T T G G T G T A C G T 2880	PHE ILE GLU PRO ASN THR ILE THR VAL TTTATTGAACCAAATAATACCATTACGGTT
--	--	--	---

FIG.210

PRO C C A 2940	ALA G C A 3000	
ARG A G A	ALA G C A	
THR A C C	GLY G G C	
THR LACA 2930	ASN 1 A A T 2990	
PHE ITT	GLY 3 G T ,	
GLU 3 A G'	SER AGTO	
ASN THR GIN ASN GLU PHE THR THR ARG PRO A A C A C A A A A T G A G T T T A C A A C C A G A C C A A 2920 2940	TTCTGAAGGTAAG 2960 2960 2970 ALA CYS PHE SER GLY ASN GLY ALA ALA GCGTGTTTCTCAAGTGGTAATGGCGCAGCA 3000	GIN C A G 3030
GIN P C A A A 2920	LY LYS G T A A 297 PHE S T T C T 0	LY GGAC.
THR A C A	ILE SER GLU GLY LYS TTTCTGAAGGTAAG 2960 2970 ALA CYS PHE SERGCGTGTTTCTCA	ALA ASP ASP GLY GIN; C T G A C G A T G G A C A G 3020 GIN ***
ASN A A C	ER GCTGO	T G A C G A 3020 GIN
	SER T T C 7 2960 AI	ASP T G A C 3020
	ILE A T	ALA G C
	THR A C A	VAL GTT
	VAL G T G	ASN A A T
	SER GIN VAL THR ILE SER GLU GLY LYS G T C A A G T G A C A A T T T C T G A A G G T A A G 2950 2970 ALA CYS PHE SER G C G T G T T C T C A A G C A A G	CYS THR ASN VAL ALA ASP ASP GLY GLN GTACCAATGTTGCTGACGATGGACAG 3010 3020 GLN ***
	SER	CYS

89/235

A G

CAG

LCDC2 hmw1A sequence

G C 180

90/235

æ

9

Ø

⊱ ARG ى G A GLY හ GTA AGC VAL GAC ں ASP ASIN Ø Ø AGA ARG C GAA Ø ALA GLU \mathcal{O} Ŋ G A G GLY GLY <u>ე</u> G ACA 現 K CTGCTTTC CCGGA GLY ... TAAC CCT PRO ...VAL ...THR GAC PR0 Ø G ပ 国 ARG G A K 国 ASN G Ø C Z, TRP \mathcal{O} \mathcal{O} <u>-</u>-⊱ ⊱ GEU K G AAA E LYS G A

园 E 贸 ...ASN

ACA

ATA

CCAACA

AGA

 \mathcal{O}

Ø

Z,

K

æ

ں

ည ဗ

AAC

G AA AAA Ø E. E. ΤÀ Ø K K A G G 160 ⊣ ں Ø A C A C 150

ASIN

LYS

開

ILE

ASN

VAL

開

ACCTTTGTTAACATAACACAAAATA 200 LYS ILE LEU VAL ASN SER ASP ILE LYSAAATCTTAGTTAATAGCGACATCAATATCAAA 220 230	GLU ASN SER HIS LEU ILE LEU TRP SER GAGAACTCCCACCTAATCCTGGAGCG 250GLU ARG ASP GLY ASN SER GLY VAL GLN ILE ASPAAAGAGATGGCAACAGCGCGTTCAGATTGAT 300	GLY ASN ILE THR SER ALA THR GLY GLY G G C A A T A T T A C T C G C T A C T G G C G A A 310 SER LEU THR VAL TYR SER SER GLY TRP VAL ASP G C T T A A C C G T T T A C T C T A G T G G C T G A T 350 360	VAL HIS LYS ASN ILE THR LEU ASN SER G T T C A T A A A A C A T T A C A C T T A A T T C A G 370
---	--	---	---

	ر	
Ċ	V	
	V	
(7	
_	=	
_		

	92/23	35	
GLY TYR LEU ASN ILE THR THR LYS SER GLY ASP GGTACTTAAACATTACGACTAAAAGTGGAGAT 390 400 400	VAL ALA PHE GLU GIN GLY ASN ASP LEU G T C G C C T T C G A A C A A G G A A T G A C C T A A 430 THR ILE THR GLY GIN GLY THR ILE THR ALA SER C C A T T A C A G G T C A A G G A A C T A T T A C C G C A A G C 450	IXS LYS GLY PHE ARG PHE ASP ASN VAL A A A A A A A G G T T T T G A T A A T G T T A 490 THR LEU SER GLY VAL LYS LYS GLY PHE LEU PHE C T C T A A G T G C A A A A A G G G G T T C T T T T T 510 540	LYS TYR SER GIN THR ASN ASN LYS A A T A C A G C C A A C A T A A T A A G 550 ASP SER ASN PHE GLU ASN HIS PHE ARG GLY THR AT A G C A A T T T C G A A A C C A T T T T A G A G G A A C T A T A G C A A T T T C G A A A C C A T T T T A G A G G A A C T 570 500

VAL GTA 620

AAA

E SER GLY
TTCAGGG7
610

E

LEU ASN TTAAATA

	93/235		
ARG 3 G C 660	TYR ' A T 720	TYR 7 A T 780	
ASN ASN A C C	SER AGTT	PRO LEU .	
ASN TRP . AACTGGA 550	ASN SA A A A A A T A	O I	
CTC	A A A	E 0	
ASIN A A 650	THR ACAA 710	GLY G G C (
GLU GAG	SER T C T	ALA 3 C C	
GLN C A G	ASN VAL SER A A T G T T T C T A	PRO GLY ALA GLY CTGGTGCCGGCC	
ARG A G G	ASN A A T (PRO	
ALA 7 3 C A A 640	 G LEU 7 I T G A 700	 SER 1	: : :
GIN C A A	ASN A A T G ARG 1	GLY 3 G C / PRO C C A T	SER ICG1
MET A T G	HIS TRP SACTGG1980 VAL THR TAACCC	SER GLY AGTGG(GLY ILE GCATAT 00
LEU MET GIN ALA ARG GIN GLU ASN TRP ASN ARGTAATGCAAGCAAGGCAGGAGAACTGGAACCGC630 640 650	CCACTGGAATG 680VAL THR ARG LEU ASN VAL SER THR ASN SER TYRTAACCCGATTGAATGTTTCTACAAATAGTTAT 720	ASN SER GLY FAACAGTGGCA 740SER ARG PRO SER PRO GLY ALA GLY PRO LEU TYRGCCGTCCATCTATAT 770	GLY ILE SER GGCATATCGT
: : :	SER T C C	ASP G A T	ASN A A T (
	ARG C G C	ILE A T T	LEU TTA?
	C G G A 670	THR ACTA O	GLY G G T
	SER T C G 67	ILE 7 A T C A 730	SER (T C G G 790
	ARG HIS SER GLY ARG SER HIS TRP ASN A G A C A C G G A C G C T C C C A C T G G A A T G 670VAL THR ARG LEUTAACCCGATT	LEU ASN ILE THR ILE ASP CTCAACATCACTATTGAT 730	ARG ARG SER GLY .GACGTTCGGGTT 790
	ARG A G A	LEU CTC?	ARG ARG SER GLY LEU ASN AGACGTTCGGGTTTAAAT 790

PCT/CA99/00938

FIG.22E

			-
	94/235		
PHE ASN ASN ASP THR VAL PHE ASN VAL ALA SERTTAACAATGACACTGTTTTAATGTTGCGTCA 820 840	GLY SER ALA VAL ASN PHE SER ILE LYS GGTTCGGCAGTTAACTTTAGCATCAAGC 850PRO PRO ILE VAL SER ASN VAL HIS ASP GLY ASNCACCAATAGTAAGCAATGTACACGACGGGAATCACCAATAGTAAGCAATGTACACGACGGGAAT870 890	HIS THR LEU PHE ASN GLY ASN VAL SER CACACATTATTCAATGGAATGTTTCAG 910VAL LEU GLY GLY ASP VAL ASN PHE HIS PHETTTTAGGGGAGGGATGTCAACTTTCATTTT930 940	ASN ALA SER SER ASN HIS TRP THR A A C G C C T C C T C C A C C A C T G G A C T C 970 HIS GLY VAL VAL ILE LYS SER GIN ASN PHE ASN A T G G C G T G G T T A T A A G T C T C A A A A C T T T A A T A T G C G T G T T A T A A A G T C T C A A A C T T T A A T 990 1020

95/23	35
-------	----

ALA SER GLU GLY SER SER LEU ARG PHE G C C T C A G G T C A G C T T A G A T T C A 1030LYS SER GLU GLY SER THR ARG THR ALA PHE THRA A A G C G A A G G T T C A A C C G C T T T T A C A 1080	ILE GLU SER ASP LEU THR LEU ASN ALA ATAGAAAGTGATTTAACTTTAAATGCCA 1090THR GLY GLY ASN ILE SER LEU ASN GLN VAL ALACTGGGGCAATATATATCATTGAACCAAGTTGCA 1140	GLY ILE ASP GLY ASN LEU GIN LYS SER GGTATTGATGGTAATCTCCAAAAAGCC 1150 LEU VAL ALA ASN LYS ASN ILE THR PHE GLU GLY TGTAGCCAATAAAACATAACCTTTGAAGGG 1100 1100 1100	GLY ASN ILE THR LEU ALA ALA ASP LYS GGCAATATCACCCTTGCAGCCGATAAA 1210 1220
---	--	---	---

(Г	7
ò	7	Į
(r	5
Ĺ	Ĺ	_

	00,4225		\$6
ASN ILE THR VAL PATATTACTGTT 1250 1260	LYS SER ALA LEU 6 2 A A A T C A G C T T T A 79 1310 1320 52	SER ALA ILE ASN TCCGCTATCAAT 1370	
LYS PRO ILE GLU ILE LYS GLY ASN ILE THR VAL AACCAATAGAAATCAAAGGTAATTACTGTT 1230 1240	LYS GLU GLY ALA ASN VAL THR LEU ARG A A A G A A G G A G C C A A T G T C A C C C T T C G T A 1270 SER ALA ASN TYR GLY ASN ASP LYS SER ALA LEU G C G C G A A T T A T G G T A A T C A G C T T T A 1320 1320	SER ILE ARG GLY ASN VAL THR ASN LYS AGTATAAGAGGAATGTCACTAATAAG 1330GLY ASN LEU THR VAL THR GLY SER ALA ILE ASNGCAATCTCACGTTACGGCTCGGTATCAAT 1350 1360 1360	ILE GLU LYS ASN LEU THR VAL GLU GLY A T A G A A A A A A A T C T T A C G T T G A A G G T A 1390SER ALA LYS PHE LEU ALA ASN PRO ASN TYR SERGTGCTAAGTTTTTAGCTAATTCCAAATTACAGC

PHE ASN VAL SER GLY LEU PHE ASP ASN TTTAACGTATCCGGCCTATTTGACAACC 1450 GIN GLY LYS SER ASN IIE SER IIE ALA LYS GLY AAGGCAAGTCAAACATTTCCATTGCCAAAGGA 1500	GLY ALA HIS PHE LYS ASP ILE ASN ASN GGGCTCACTTTAAAGACATTAATAACA 1510THR LYS SER LEU ASN ILE THR THR ASN SER ASPCTAAGAGTTTAAACATTACTACCAACTCCGAC1530 1540 1550	SER ALA TYR ARG THR ILE ILE GLU GLY TCCGCTTACCGCACTATTATAGAAGGCA 1570ASN ILE THR ASN SER ASN GLY ASP LEU ASN ILEASN ILE THR ASN SER ASN GLY ASP LEU ASN ILEATATAACCAACAGTAACGGGATTTAAATATC1590 1600	THR ASP ASN ASN ALA GLU ILE ACTGATAATAAAATAACGCTGAAATCC 1630
--	---	--	--

SUBSTITUTE SHEET (RULE 26)

98/	235
-----	-----

	98	/235		
GLN ILE GLY GLY ASN ILE SER GLN LYS GLU GLY A A A T T G G C G C A A T A T C T C G C A A A A G G G G T1650 1660 1660 1670 1680	ASN LEU THR ILE SER SER ASP LYS ILE A A T C T C C C A T A A A T T A 1690 ASN ILE THR ASN GIN ILE THR ILE LYS LYS GLY	A T C A C T A A C C A G A T A A C A A T C A A G A G G T1710 1720 1730 1740	VAL ASN LYS GLU ASP SER SER SER GTTAATAAGGGATTCTGATTCAAGCA .1750	THR ALA ASN ASN ALA ASN LEU THR ILE LYS THRCGGCAAACAATGCTAATCTAACCATTAAAACC 1770 1780 1780

A C G G G A G A C C T A A... 1820 ASP GLY]] T T A EEI TGCAA 1810 GIN EEG. ٤٦ GAA GEU AAA

A C A 1860 ATC A A A G C A G A A 1850 TCAGGCTTCGAT ASP GLY..ASN

... A T A

GEU

LYS

ALA

99/235

G T A1980 G G C VAL ⊱ AAT ASIN A A A A CACi ALA AAT ပ ဗ ASIN AACAAT ⊱ ASIN GAAAGC 段 SER GACAG ASN ASP GU AAC $G \subset T$ AATGACGCT ASN ALA ALA GATAAT ATCTCT ASIN SER ASP 1900 AACCAGGTTAAAG... ASIN ACCTCTA... AGT ... A T T C G A A A AAT GCA LYS ASN J L ... GTAAT ILE ASIN AAT .. A T G G C SER 開 GLY ASN Ø GATTTA GTAGAA ...1890 LEG · · · GLY ...ASP GLN ...1950 ACTATT . ASN ...2010 ILE 1940 ASP ASIN VAL 强 $C \subset C$ TLL ALA AAA ТТА LYS 园 E TAACC ACAGT GGCACCAGC GAGGG SER SE WAL ASN G Ø AAA A A K 园 Ø ى AA CA GAT

\mathbf{X}	
2	
2	
Ġ	
正	

	100/235		
SER T C T 2100	ILE A T T 2160	VAL G T T 2220	ASN A A C 2280
THRACTI	THR	LYS A A A (SER A G C 7
ILE A T T A	THR ACA	LYS GLY AAAGGTP	ALA VAL ; C T G T A P
ASN ASN AACAATA 2090	ALA GLY THR SCGGCACAP 2150	LYS L A A A 2210	ALA P G C T (2270
ASN A A C	ALA G C G	ILE ATTA 22	LEU CTTG 22
LYS ASN ILE THR VAL ASN ASN ASN ILE THR SER. AAAATATAACAGTAAACAACAATATTACTTCT 20070 20090 2100	THR ALA SER CTGCGTCAG 20 310 ASN VAL THR LYS ALA GLY THR THR ILE A A A A T G T T A C C A C C A A G C G G C A C A T T T C T A C C A C A A G C G G C A C A T T T A C C A C C A A G C G G C A C A T T T A C C A C C A A G C G G C A C A T T T A C C A C C A A A G C G G C A C A T T T A C C A C C A A A G C G G C A C A T T T A C C A C C A A A A A T G T T A C C A C C A A A G C G G C A C A A C C A T T T A C C A C C A A A A	VAL GLU VAL 3 T A G A A G T A A 180 17HR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL 17TR ALA LYS THR GLY ASP ILE LYS GLY LYS VAL	VAL THR LEU 3 TAACACTTA 240 THR ALA THR GLY GLU ALA LEU ALA VAL SER ASN CTGCAACCGGAGAGCTCTTGCTGTAAGCAAC 2250 2250
IR VAL 2 A G T A 2080	THR ALA SER ACTGCGTCAG 120 GLU ASN VAL THR THR .AAAATGTTACCACCA	GLU VAL 3 A A G T A A ALA LYS THR GLY 3 C C A A A C A G G T C	 GLY GLU 3 G A G A A C 2260
THR ACA 20	 THR 7 ACCA 2140	 THR C A C A G	 A GLY (G G A G 2260
ILE 'ATA	THR ALA SER A C T G C G T C A G 120 GLU ASN VAL THR . A A A A T G T T A C (2) 2130	VAL GLU VAL 3 T A G A A G T A A 180 1THR ALA LYS THR . C A G C C A A A A C C	GLU SER THR SER GLY SER VAL THR LEU GAATCCACTTCCGGCTCTGTAACACTTA 2230 2240THR ALA THR GLYCTGCAACCGG?
ASIN A A T	AIA G C G ASN A A T	VAL GLU G T A G A A C 180 THR ALA C C A G C C A	VAL THR LEUSTA ALA CACTA 240 THR ALA THR CACA ACCA 2250
LYS A A 2070	A 11	VAL C G T A (2180 THR C A C	VAL T G T A 7 2240 THR C T G
• • •	ILE 'ATC' '	SER A G C	SER TCT
	ASN A A T A	THR GLY CAGGT,	SER GLY
	IR VAL 2 A G T A A 2110	THR ACA 70	SER TCC 30
	THR A C A C 2110	THR 3 ACCA 2170	THR SACTT 2230
	HIS LYS THR VAL ASN ILE CACAAAACAGTAAATATC 2110	ASN ALA THR THR GLY SER A A T G C A A C C A C A G G T A G C (2) 2170 2	GLU SER THR AATCCACTT 2230
	HIS C A C A	ASN A A T	GLU GAA

FIG.22L

	101/235		
GLY G G C 2340	GLY G G T 2400	GLY G G C 2460	
ALA G C A O	SER AGC	SER T C C	
THR GIN ACTCAAC 330	ILE A T T	LYS A A A T	
THR A A C T 2330	ASP C G A T P 23390	THR TACT 7 2450	
THR ACAA 23	GLY GGCC	THR ACTP	
TACCATCACTG 2300 ALA ASN LYS LEU THR THR GIN ALA GLY CAAATAAGGGTAAATTAACAACTCAAGCAGGC 2340 2320 2320 2320	TAACGGTGTAA 2360 THR ALA SER GIN SER GLY ASP ILE SER GLY CTGCCTCAATCAGGCGATATTAGCGGT 2370 2380 2400	AGTAAAAGTTA 2420 SER ALA ILE GLY ASP LEU THR THR LYS SER GLYGTGCGATCGGTGATTTGACTAAATCCGGC2430 2440 2450	
LYS A A A T	GIN CAA7	ASP GATT	
 GLY I G G T A 2320	 A SER C A G C C 2380	LYS VAL A A G T T A ALA ILE GLY A C G A T C G G T G	: : :
TACCATCACTG 2300ALA ASN LYS GLYCAAATAAGGG'	TAACGGTGTAA 2360THR ALA SER SERCTGCCTCAAG	VAL GTT ILE ATC	GLU GAG
THR ILE CCATC 00 ALA ASN CAAT CAAAT 2310	ASN GLY A A C G G T S60 THR ALA C T G C C 2370	VAL LYS TAAAA 20 SER ALA GTGCG	THR GLY CAGGT 80
TACCATC 2300 ALA ASN CAAAT	T A A C G G T 2360THR ALA C T G C C C 2370	A G T A Z 2420 SER G T G	THR AACA 2480
		ASN THR ACACA	ALA LYS
ASN THR VALA A CACTGT	ALA G C G	ASN A A C	ALA G C A
JY ASN 3 C A A C 2290	AL SER FTAGCG 2350	R GLY CGGTA 2410	LYS A A G G 70
B)		# C	SER GLU ILE LYS ALA LYS THR GLY GLU TCGGAAATCAAGGCAAAAACAGGTGAGG
SER	THR ACGG	ILE SATT	SER GLU CGGAAA
ILE ATTT	SER TCTA	THR A C G A	SER T C G

FIG.22N

	102/235·		
GLY G G T 2520	GLU G A A 2580	LYS A A A 2640	ALA G C T
ILE A T T C	VAL G T T	C G C	SER T C A
GLY THR 3 G T A C A A 310	TAR ACT	SER T C G (LEU
GLY A G G T 2510	ASP LEU THR VAL GLU FATTTAACTGTTGA 2570 2580	THR AACAT 2630	ASN A A C (
THR ACAG 25	ASP GAT	ALA ' G C A A 26	VAL 3 T A ,
ALA G C G A	THR GLY	THRACT	GIN C A G (
ALA ASN VAL THR SER ALA THR GLY THR ILE GLY C T A A C G A C A G T G C G A C A G G T A C A A T T G G T2490 2520	AIA VAL ASN TGCAGTAAATG 2540VAL THR ALA ASN THR GLY ASP LEU THR VAL GLUTACAGCAAATACTGGCGATTTAACTGAA 2550 2550	AIA THR GLY TGCGACAGAG 2600GLY ALA ALA THR LEU THR ALA THR SER GLY LYSGAGCGGGCTAACTGCAACATCGGGCAAA 2620 2640	SER SER ILE LTCAAGCATTA 2660 THR SER ALA ASN GIN VAL ASN LEU SER ALATHR SER ALA ASN ASN GIN VAL ASN LEU SER ALA
THR S A C A A 2500	AIA VAL ASN T G C A G T A A A T G 2540VAL THR ALA ASN 7T T A C A G C A A A T A T T A C A G C A A A T A2550 2560	 G THR I A C C C	 ASN A A T
VAL G T G	ALA VAL ASN TGCAGTAAATG 2540VAL THR ALA ASNTTACAGCAAA	THR GLY CAGGA ALA ALA CCGCG	SER ILE AGCATT SER ALA CAGCT
ASN A A C	ALA VAL ASN CAGTAAA 40 VAL THR ALA TTACAGCA	THR ACA ALA GCC	SER AGC SER TCA
AIA C T 2490	T G C A 2540 VAL T T T 2550	T G C G A 2600GLY G A G	SER T T C A A 2660 THR S
• • •	-	ASP G A T	SER AGT
	SER GLY	ILE A T T	ALA G C T ?
	SER T C T O	LYS A A A i O	LYS A A G G O
	ILE :	ALA ALA 1 CCGCAA 2590	THR THR I.CCACTA 2650
	SLY THR ILE SER GLY ASN GTACGATTTCTGGTAA 2530	ATGCCGCAAAATTGATGCGACAGGAG 2590 2600GLY ALA ALA THRGLY ALA ALA THRGLY ALA ALA THR 22610 28	EU THR LYS ALA SER SER ILE TAACCACTAAGCTAGTTCAAGCATT 2650 2650THR SER ALACTTCAGCT
	SLY G T	ASP A T G	JEU TA 1

G G C A G C G T A 2900

AATGGC1 2890

ASIN

FIG.22N

	INS ASP GIY SER IIE GLY GIY ASN IIE A A G G A T G G T A G C A T T G G G G A A T A T C A 2710 ASN AIA AIA ASN VAL THR LEU ASN ASN AIA ASN VAL THR LEU ASN 2730 AIA LEU THR THR VAL LYS GLY SER SER 2770 ILE ASN AIA ASN SER GLY THR LEU TT A A C G C A A C G C A C C T G TT A A C G C A A C G C C C T G TT A A C G C A A C G G C A C C T G AIA LYS ASP AIA GLU LEU ASN GLY GLU 2830 AIA SER GLY HR VAL VAL VAL VAL VAL VAL VAL ASN TA A C G C A A C G G C A C C T G TT A A C G C A A C G G C A C C T G TT A A C G C A A C A G G C A C C T G TT A A C G C A A C A G G C A C C T G TT A A C G C A A C A G G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C A G C G C A C C T G TT A A C G C A A A C G C G C A C C T G TT A A C G C A A A C G C G C A C C T G TT A A C G C A A A C G C G C A C C T G TT A C C C A A A C G C C A A C G C G C A C C T G TT A C C C A A A C G C C A A C G C G C A C C T G TT A C C C T G A C C T A A C G C G C A C C T G TT A C C C T G A A C C C T G C C C T G C C C T G C C C T G C C C T G C C C T G C C C T G C C C T G C C C T G C C C T G C C C T G C C T G C C T G C C C T G C C T G C C C T G C C T G C C C T G C C T G C C T G C C T G C C T G C C T G C C T G C C T G C C T G C C T G C C T G C C T G C C C T G C C C T G C C T G C C T G C C T G C C T G C C T G C C T G C C T G C C C T G C C C T G C C C T G C C C T G C C C T G C C T G C C C T G C C C T G C C C T G C C C T G C C C T G C
--	---

SUBSTITUTE SHEET (RULE 26)

103/235

ATTAAC 2820

' A C A G G C 2760

(C
Č	V
C	V
(5
Ī	

	104/235		
GLY G G G 2940	LEU C T G 3000	GLU G A A 3060	GLU 3 A A 3120
THR ACTG	VAL G T A C	ASP GAT	ARG A G A (
ILE A T C A	THR	VAL G T A	GLU 3 A A A
ASN 3 A A C 2930	ILE ASN THR TAAACACCO 2990	ALA SER VAL 3 C A A G C G T A C 3050	GLU GLU FGAAGAA 3110
ARG VAL ASN IGAGTGAACA 2930	ILE ATA 2	ALA G C A	ASP GATG
.ALA THR THR SER SER ARG VAL ASN ILE THR GLY . CGACAACTCAAGCAGAGTGAACATCACTGGG .2910 2920 2930 2940	GLY LEU ASN GGATTAAATA 1960 ILE ILE SER LYS ASN GLY ILE ASN THR VAL LEU .TCATTTCAAAAACGGTATAAACACGTACTG	ASP VAL LYS GATGTGAAAT 3020 TYR ILE GIN PRO GLY ILE ALA SER VAL ASP GLUACATTCAACCGGTATGAA 3040 3050	ILE LEU GLU A T C C T T G A G A 3080 LYS VAL LYS ASP LEU SER ASP GLU GLU ARG GLU . A G G T A A A A G A T T T A T C T G A T G A A G A A G A A G A A A G A G A A G A G A A G A G A A G A G A A G A G A A G A G A A G A G A A G A G A G A A G A G A A G A G A G A A G A G A G A A G A G A G A A G A G A G A A G A G A A G A G A G A A G A G A G A G A G A A G A
SER AGCA 20	ASN A A C 30	 PRO GLY . C G G T A	 ASP LEU 3100
SER S T C A A 2920	ASN A A T A SER LYS ASN C A A A A A A C C	 T PRO C C C G G	 ASP I G A T T 3100
THR	GLY LEU ASN G G A T T A A A T A 2960 ILE ILE SER LYS T C A T T T C A A A	ASP VAL LYS G A T G T G A A A T 3020TYR ILE GIN PROA C A T T C A A C C C	ILE LEU GLU A T C C T T G A G A 3080 .LYS VAL LYS ASP . A G G T A A A A G A'. 3090
THR ACAA	LEU ATTAI ILE ATTI	VAL GTG1 ILE ATTC	ILE LEU T C C T T C 180 LYS VAL A G G T A A 3090
ALA C G A 2910		ASP G A T G 3020 TYR A C A	ILE 3080 LYS A G G
	ASN A T A A T	ILE ATT	ARG C G C
	ILE A T A .	LYS A A A A	LYS A A A
	LEU ILE THR PAATCACAA 2950	VAL G T T A 0	ALA G C G O
	ILE A T C A 2950	LYS GLY VAAGGCG3010	GLU <i>P</i> 3 A A G 3070
	LEU T T A	LYS A A A	VAL ILE GLU ALA LYS ARG TAATTGAAGCGAAACG
	ASP LEU ILE THR ILE ASN GATTAAT 2950	LEU LYS GLY VAL LYS ILE TTAAAAGGCGTTAAAATT 3010	VAL ILE GLU ALA LYS ARG GTAATGAGGOGGCGAAACGC

1 A T T G C T (3290

ILE

ASIN

温

FIG.22F

			VAL ARG PHE ALA GLU PRO ASN ASN ALA	TACGTTTTGCTGAGCCAAATAATGC	3170			ARG PRO SER SER GLN VAL THR ILE SER	G A C C A T C A A G T C A A G T G A C A A T T T C 1	3230		
			GLU PR	GAGCC	0			GLN VA	CAAGT	0		
•	.: G:	:	ALA	GCT	3160	:	. :	SAS	AGT	3220	:	.: G.
LEU ALA LYS LEU GLY VAL SER ALA	GCGTTAGCTAAACTTGGCGTAAGCGCTG		田田	TTT		LE ASN THR GIN ASN GLU PHE THR THR	АТТААТАСАСААААТGАGТТТАСААССА 3190	段	$T \in A$		LYS VAL CYS PHE LEU ILE GLY ASN GLY	TTTCTTAATCGGCAATGGTG
SER	AGC		ARG	CGT		TH.	ACA	PRO	CCA		ASIN	AAT
VAL	GTA	3140	VAL	T A	3150	田田	3 T T T 3200	ARG	G A	3210	CLY	2 G G C
GLY	0 9 g		•	•	•	CIN	G A G	•	•	•	ILE	ATC
EEG	C T T					ASIN	A A T				LEU	ΤΤΑ
LYS	AAA	3130				GLN	3190 A A				PHR	TTTC
ALA	GCT	31				强	A C A 31				CYS	
LEG	ΤТА					ASIN	A A T				VAL	AAGGTATG
ALA	909					ILE	ATT				LYS	AAG

105/235

ARG ** CGGTA

GLY G G G 180

160

ICDC2 hmw2A sequence

	()	2	<u> </u>
	ARG C G C	LYS A A A	ARG A G G
	GLU 3 A A	GLN C A A	LYS
	ALA 3 C T (HIS C A T (LEU TAP
	SER CG(ASN . A C C	MET T G T
	PRO CTI	ILE TTA	SLY 1
	CCCGGATGATG 20VAL SER ILE ASP ALA PRO SER ALA GLU ARGTATCCATTGACGCACCTTCGGCTGAACGCA 30 40 50	GGAATACACCG GGAATACACCG 80 GLY THR GLY ALA ASP ILE ASN HIS GIN LYSGAACAGGGCTGATATTAACCATCAAAAAC90	AACATTAACAA 140ASN THR THR LEU GLU GLY MET LEU LYS ARGACACAACTCTTGAGGGGATGTTAAAAGGG
	3 ASP 7 ACG	 3 ALA 7 1 C T G	 A LEU (CTTG
	LEU ASP PRO ASP ASP T T A G A C C C G G A T G A T G 20VAL SER ILE ASPT A T C C A T T G A (2.1)	ASP VAL GLU TYR THR GACGTGGAATACACCG 80GLY THR GLY ALAGAACAGGGCC	THR ACAP THR CCTC
	PRO ASP ASP CGGATGAT 20 VAL SER ILE TATCCATT 30	GLU TYR THR 3 A A T A C A C C 80 GLY THR GLY G A A C A G G G	THR LEU THR ACATTAACA 140 ASN THR THR ACACAACT
-	C C C G C 20 VAL T A 5	GLU G A A C 80 GLY G A A	THR A C A 7 140 .ASN
	LEU ASP TAGACO	VAL 3 T G G 	LYS SER AAGTCAA 1
	LEU T T A	ASP 3 A C G	LYS A A G C
			THR A C C Z
	TRP T G G T 10	GLY 3 G C (GLU 7 3 A A A 130
	LYS GLU AAGAGT	ASP THR GLY GLU ACACTGGCGA7	ASN SER GLU THR ACAGCGAAACO
	LYS GLU TRP LEU AAAGAGTGGTTG 10	ASP THR GLY GLU GACACTGGCGAA 70	ASN SER GLU THR LYS SER THR LEU THR A A C A G C G A A C C A A G T C A A C A T T A A C A A 130ASN THR THR LEU

SUBSTITUTE SHEET (RULE 26)

106/235

GLN C A A 120

ACGCACT S

FIG.23E

	107/235		
CTTTTTTTTTTTATCACCGCCAGAATA 190 LYS ILE ARG VAL ASN SER THR ILE ASN ILE GLY A A A T C C G A G T T A A T A G C A C C A T C A T C G G G 210 220	ASP SER GLY HIS LEU THR LEU TYR LYS GATAGCGCCATTTAACCTTTACAAA 250 LYS ARG LYS ASN ARG SER ASP GLY ILE GLN ILE AAAGAAAAATCGTAGCGATGGTATTCAAATT AAAGAAAAATCGTAGCGATGGTATTCAAATT 280 280	ASN LYS ASP ILE THR SER THR GLY GLY A A C A G G A T A T T A C T T C T A C G G C G G A A 310SER LEU THR ILE ASN SER ASP ASP TRP VAL ASPGTTTAACTATTAACTCCGACGACTGGGTTGAT 360	ILE HIS GLY ASN ILE THR LEU GLY GLU A T T C A T G G A A T A T C A C G C T T G G T G A G G 370

SER	TCC	
ASP	G A T	
SER	A G T	410
SER	$T \subset T$	
開	ACC	
TIE	ATT	400
ASIN	TTTTAAATATTACCTCTAGT	₫(
LEU	ТТА	
出	T T T	
\cdots GLY	G C T	390

			ALA	909	480
			ILE	ATC	
			ILE	ATT	
			GEN	CAA	470
			SER ALA GLN	GACGTAGCTCAGCAAGTGCTCAAATTATCGCG	
			SE	AGT	
			ALA	$G \subset A$	0
:	 G.:	:	S	TCA	460
LYS	AAA		SER	A G C	
ASN	AAT		ARG	CGT	
ASN GLY	ACGGCAATAAAG	440	.GLY	.GA	. 450
ASN	A A ·C		•	:	:
GLY	G G A				
GLY	G G T	0			
CIN	GAG	430			
出	TTC				
ALA	GCTTTCGAGGTGGA				

108/235

			SER	I C T	
			VAL	. A C C T T T A G A C T C A A C A A T G T G T C T	
			ASIN	A A T	530
			ASN ASN VAL	AAC	
			LEU	CTC	
			ARG	A G A	0
:	. A	:	思	T T T	520
CITO	CTTACTGGAGAA		LYS THR PHE	ACC	
GLY	G G A		LYS	ATAAA	
開	$\mathbb{A} \subset \mathbb{T}$	200	ASIN	. A T	510
LEU THR	C T T		:	:	•
置	$A \subset T$				
HE.	ATA	90			
国	CAGGGTACT	490			
GLY	G G T				
GIN	CAG				

LEU TTA 540

			ARG	A G A	
			HIS ARG	CAT	
			段	TTTCAACAGCAAGCAATTTATCTCATAGA	290
			LEG	TTA	
			ASIN	AAT	
			SER	AGC	0
:	A	:	ALA	GCA	580
H	AATGGTCTAAGTATTA.		出	ACA	
SE	A G T		SER	T C A	
	C T A	260	ILE.	₽	. 570
GLY	GGT		•	•	:
ASIN	AAT				
GLY	G	00			
出	3 G G A C G G	550			
GLY	G G				

-1G.23D

ASP GLY GLU ILE ASN VAL SER GLY ASN G A C G G T G A A T G T A T C T G G A A A T G 610VAL THR ILE ASN GLN THR GLN GLN ASN ILETAACAATTAATCAAACCACGCAGCAAAACATT 650 660	GLU TYR TRP LYS ALA SER SER ASP SER GAATACTGGAAGGCTAGCGATTCTT 670TYR TRP ASN VAL THR SER PHE ASN LEU ARG GLUATTGGAATGTCACTTTTAATTTGAGAA 720	ASP SER LYS PHE THR PHE ILE LYS TYR GATTCAAAGTTTACCTTTATCAAATACG 740 VAL ASN SER ALA ARG ASN GLY ASP VAL ARG GLY VAL ASN SER ALA ARG ASN GLY ASP VAL ARG GLA TTAACTCTGCCAGAAATGGTGATGTAAGAGAA 770	ARG SER PHE ALA GLY VAL ILE PHE ASN A G A A G T T T G C A G G T G T G T T T T A A T G 790 800
--	---	---	---

L	Ļ	3
C	Υ)
Č	\	1
r	Γ	;
<u>'</u>	_	_
L	1	_

	110 /225		
/D 🔾	110/235		
ALA LYS GLY LEU THR THR SER PHE ASN VAL LYS CTAAAGGTCTCACTACAAGCTTTAACGTCAAG	TTTTAAATTAA 860LYS PRO ASN SER GLY TYR ASN SER GLN LYS ARGAGCCAAATTCAGGCTATAATTCACAAAAAGGAGCCAAATTCAGGCTATAATTCACAAAAAGG	SER ASN ILE T C C A A C A T C T 920 SER VAL SER GLY GLY ARG VAL ASN ILE ASN C G G T C T C A G G A G G A A G G G T A A A C A T T A A C 950	SER T C T 1020
VAL G T C	LYS A A A	ILE	VAL ; T T ;
PHE ASN VAL TTAACGTC1 30	GIN	ASN A C Z	ASN A A T G
PHE 7 7 7 7 330	SER GLN LYS TCACAAAAA 890	VAL ASN GTAAACA	ILE , 1 T A 10
SER AGCI 8	ASN ASIV 8.8	ARG 1 1 G G G	SER ILE CAAT 1010
THR	TYR P	GLY A	SER SI
TH LAC	TY.	GE G	SE A G
THR 'ACTA	. GLY A G G C T 880	GLY G G A	SER T C G A 0
LEU CTC	 A SER T C A G	 T GLY (G G A G	 G ARG S A G A T 1000
GLY 3 G T (LEU TTA ASN A A T	ILE A T C SER	GLY 3 G A (ILE 1 T A P
LYS A A A G	LYS A A A T PRO	ASN A A C ? VAL 3 T C T	GLY GGAG GLU GAGA
. C T A	PHE LYS LEU T T T A A A T T A A 860LYS PRO ASN SER A G C C A A A T T C A	SER ASN ILE T C C A A C A T C 920 SER VAL SER C G G T C T C A 930	GLY GLY GLY GGCGGAGGAG 980 VAL GLU ILE ARG SER SER SER ILE ASN VAL SER .TTGAGATAAGATCGAGTTCAATTAATGTTTCT 1020
	ASP 3 A T T 8 8 8		000
	VAL ASTIG	PHE C	LEU T
	THR VA C A G '	A T A	C E
	3 1日日 G A C 850	3 GLN T C A A T 910	ASN A A A V
	SER . T C G A 850	ILE ATT 91	ALA ASN 3 C C A A T C 970
	LIS GLY SER THR VAL ASPAAAAAG TIGAT 850	ILE PRO ILE GIN PHE GIN ATTCCAA 910	THR LEU ALA ASN LEU THRACGCCAATCTTACA 970
	L.7S A. A. G.	ILE TTC	71张 , C G C
	V	A	A

E E CI

ASP GLY G A T G G C

	111/235		
THR A C C 1080	GLY 3 G A 1140	GLY 5 G T 1200	
ATGACAGCTC 1040 GIN ALA ARG ASP ARG ASN ALA PHE GLU ILE THR AGGCTCGCGACAGGAATGCCTTTGAAATTACC 1050 1050 1060 1060	ALA SER ASN GCAAGCAATT 1100 SER ASN LEU SER ILE GIN GIN ASN ASP GLYCAAACCTATCTATTACAGCAAAATGATGGA	ALA ASN ALA GCAAATGCCA 1160 ILE ASN SER LYS TYR ASN VAL THR ILE GIN GLY .TTAACTCAAATATAACGTAACTATTCAAGGT .1170 .1170	
U. J. A.	NA T G		
G T G A	A A A	II A T	
A PHE CTTT 1070	GIN 3 C A A 1130	THR A A C T 1190	
ALA G C C	ILE GIN GIN ASN TACAGCAAAAT 1130	VAL GTA	
ASN A A T	ILE A T A	ASN 1 A C	
ARG ASN ALA PHE GLU AGGAATGCCTTTGAAA 1070	ILE TT?	IYR A T <i>P</i>	
ALA C T C ARG ASP ACG C G A C A 1060	ALA SER ASN G C A A G C A A T T 1100 SER ASN LEU SER ILE C A A A C C T A T C T A T T A 1110	 LYS 1 A A T 1180	: : :
	NT T T T T T T T T T T T T T T T T T T	A C A R I A A A	
AG C AR AR I C G	AS C A P LE C C T	AI TGC SE	AS A A A
ACAC ALA GCTC	ALA SER ASN CAAGCAA7 00 SER ASN LEU CAAACCTA	ALA ASN ALA ; C A A A T G C C 60 ILE ASN SER T T A A C T C A 1170	GLY GLN ASN GGCAAAAT 220
A G G C A G G C A G G C A G C	ALA 1100 SER C A	ALA ; G C A 1160 ILE T T	GLY G G G 1220
	ASN A A C. 1	LYS A A G 1 1	GLY 3 G C
D E	LYS ASP LEU VAL ILE ASN ALA SER ASN A A A G A T T T A G T T A T A A A C G C A A G C A A T T 1090SER ASN LEU SERCAAACCTATC'	ASP ASN ASN GIN ATAATAATCAAA 1150	GLY ASN VAL THR LEU GLY GTAATGTTACCCTTGGO 1210
S C C C	VAL 3 T T A	ASN A T C	展 C C C
2 A A 1030	TU (N	NL 7 F T A 1210
7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		A T A P	VP TGT
0 0 0	LYS ASP LEU AAGATTTAG 1090	ASP G A	ASN A A A
GATGGCTCAACCCTCTATGACAGCTC 1030 GIN ALA ARG ASP 1050 1050 1070 1070 1070 1070 1070 1070	LYS A A A	PHE ASP ASN ASN GIN LYS ALA ASN ALA TTTGATAATCAAAAGGCAAATGCCA 1150ILE ASN SER LYSTTAACTCAAA.	GLY ASN VAL THR LEU GLY GLY GLN ASN GGTAATGTTACCCTTGGCGGGCAAAATT 1210

C	Г)
C	Υ,)
C	1	1
_	r	
(÷	J
Ī	ī	_

	112/235		
ILE A T T 1260	LEU C T A 1320	ILE A T T 1380	LEU C T A 1440
SER SER THR ILE THR GLY SER VAL ASN ILE CAAGCAGTACAATCACAGGGAGTGTTAATAT 1230 1240 1250	THR LEU GLN CTTTGCAAG 180 ALA HIS ASN GLY ASN ASP ARG ASN LYS LEU CCCACAATGCAATGATAGAAAAAAGCTA 1300 1300 1300	VAL GLU GLY 3 T T G A A G G A G 140 GLU LEU ARG LEU VAL GLY ALA SER ALA ASN ILE A A T T A A G C T A G T T G G C G C A A G T G C A A T T A 1360 1350	LYS SER GLY A A G A G C G G T G 100 ALA LYS PHE LYS ALA GLU THR ASN ASP ASN LEU C T A A A T C A A G C A G A A C A A T G A C C T A 1410 1410
VAL G T T A	LYS A A A A	ALA G C A 2	ASP , G A C A
SER 3 A G T C 1250	3 ASN A A A T A 1310	SER A A G T C 1370	ASN A A A T G 1430
G G G	ARG A G A A 13	ALA G C A A 13	THR ACAA 14
SER THR ILE THR GTACAATCACA(ASP GAT?	GLY G G C C	ALA GLU SCAGAA?
IR ILE CAATC 1240	GLN C A A G ASN GLY ASN A A T G G C A A T C 1300	 JU VAL PAGTT 1360	 .YS ALA A A G C A 1420
THR TACA	 GLY 13	GLY 3 G A G ARG LEU 1360	SER GLY AGCGGTG LYS PHE LYS AATTCAAAG
SER A G T	THR LEU GIN CTTTGCA 80 ALA HIS ASN CCCACAA 1290	GLY AGG 7 ARG ARG	LYS SER GLY A A G A G C G G T 1400 ALA LYS PHE C T A A A T T C .1410
SER AGCA	THR LEU A C T T T G (280 ALA HIS C C C A C 7 1290	VAL GLU GLY 3 T T G A A G G 7 440 GLU LEU ARG A A T T A A G C	LYS SER GLY A A G A G C G G G 400 ALA LYS PHE C T A A A T T C
C A A	THR T A C T 1280 ALA C C	THR PHE GLY ASN VAL SER VAL GLU GLY A C C T T C G T A A T G T A T C T G T T G A G G A G 1330GLU LEU ARG LEUA A T T A A G C C T.	7
• • •	VAL GTT	SER TCT	VAL GTT
	ASN A A T	VAL G T A I	SER AGT
	ALA G C A 10	ASN A A T	LEU CTT
	ASN A A A T G 1270	GLY 7 G G T A 1330	ASN I A A T C 1390
	GLY ALA ASN ALA ASN GCGCTAATGCAAATG	THR PHE GLY ASN CCTTCGGTAATG 1330	ASN ASN LEU SER VALA A CAACAATCTTAGTGTT 1390
	GLY ALA ASN ALA ASN VAL THR LEU GLN GGCGCTAATGCAAATGTTACTTTGCAAG 1280ALA HIS ASN GLYCCCACAATGG	THR A C C	ASN A A C

ASIN

ASIN

開

開

GLY

强

ILE

ASN

	ALA G C G	1500		ASN A A C	9		THR A C G	1620	
	GCACCTCCATAATTGATGTAAAAAAGGGGCG			GLY ASN LEU ASN ILE THR THR ASN ALA LYS ASNGTAATTTAAATATCACTACTAATGCTAAAAC			ILE THR ASN ASN LYS GLY ALA LEU ASN ILE THR		
	ASP VAL LYS LYS			THR ASN ALA CTAATGCT?			ASIN AATA		,
	LYS A A A	1490		ASN A A T			ALA LEU	1610	
	VAL GTA			THRACT	•		ALA G C T	•	
	ASP G A T			ILE THR TCACTA			GLY G G T		
	GCACCTCCATAATTG	1480		ILE AT C) [THR ASN ASN LYS GLY	00	
.: .:	ILE	14		ASN A A A T A	•	Г А	ASN S A T	1600	: : : : : : : : : : : : : : : : : : :
C A A	SER CTC		ASP TGA	LEU	GLY ASN	AAA	ASN FAA(GLN C C A 7
CAA	JACO	0	C A A '	ASN TAA T		7 9 9 D	THR		ILE A A T (
TTACCAACAACG 1460	GLY	1470	ILE THR ASN ASP TTACCAATGATG 1520	GLY ASN LEU GTAATTTAA	ASN	TCAACGGAAATA 1580	ILE	1590	THR GLU ILE GIN ACTGAAATCCAAA 1640
C I I			I ILE TAT		Ħ	TAT			
CAC			CAAT		SER VAL	C G T			J ASP TGAC
C G G 1450			M GLY PAGGC 1 1510			A A G 1570			GLY ASN 3 G T A A T G 1630
TAC			S LE		GIN LYS	AAA			ASN GLA
AACATTACCGGCACCT 1450			ALA LYS LEU GLY ASN CCAAAACTAGGCAATA		CLY GL	GGTCAAAAAAGCGTTA 1570			ASN ASN GLY ASN ASP ? A A T A A T G G T A A T G A C A 1630
A A			AI G C		덩	9			AS A A

_	_	_
C	~)
Ć	`	J
		•
ſ	Г	5
`	Ξ	_
L	i	

	114/235		
ILE GLY GLY ASN ILE SER GLN LYS GLU GLY ASN T T G G C G G C A T A T C T C G C A A A A A G G T A A T A T C T C G C A A A A G G G T A A T A T C T C G C A A A A G G G T A A T A T C T C G C A A A A G G G T A A T A T C T C G C A A A A G G G T A A T A T C T C G C A A A A G G G T A A T A T C T C G C A A A A G G G T A A T A T C T C G C A A A A G A A G G T A A T A T C T C G C A A A A G A A G G T A A T A T C T C G C A A A A G A A G G T A A T A T C T C G C A A A A G A A G G T A A T A T C T C G C A A A A A G A A G G T A A T A T C T C G C A A A A A G A A G G T A A T A T C T C G C A A A A A G A A G G T A A T A T A T C T C G C A A A A A G A A G G T A A T A T A T C T C G C A A A A A G A A G G T A A T A T A T C T C G C A A A A A G A A G G T A A T A T A T C T C G C A A A A A G A A G G T A A T A T A T C T C G C A A A A A G A A G G T A A T A T A T C T C G C A A A A A G A A G G T A A T A T A T C T C G C A A A A A G A A G G T A A A A A A A A A	LEU THR ILE SER SER ASP LYS ILE ASN C T C A C A A A T T A A T A 1700 ILE THR LYS ARG ILE GLU ILE LYS ALA GLY THR T C A C C A A A C G G A T A G A A T T A A G G C A G G T A C T 1740 1740	ASP GIN GLY ASN SER ASP SER GLY VAL GATCAAGGAATTCTGATTCAGGCGTAG 1760 ALA SER ASN ALA ASN LEU THR ILE LYS THR LYS CAAGTAATGCTAATCTAACCATTAAAACCAAA 1800	GLU LEU LYS LEU THR GLU ASN LEU ASN GAATTGAAATTAACAGAAACCTAAATA 1820ILE SER GLY PHE ASP LYS ALA GLU ILE VAL ALATTTCAGGTTTTGATAAAGCAGAAATTGTAGCC1830 1840 1850

FIG.23.

II.E

ASN A A T

ASIN

ASN

LEUTA

AATAAC*P* 1870

GLU GAG

	115/235					
THR A C A 1920		LEU 7 T A 1980		ASN A A C 2040		
LYS A A A A		THR ACAO		ASN ATA		
ALA 3 C C A		VAL 3 T G A		ASP 3 A C A		
ASN T A A C (1910		ASN 2 A A T C 1970		SER 1 A G T C 2030		
ALA G C T 7		HIS CACI		ASN A A T A		
ASIN A A T		GLY GGT		GLY 3 G T /		
ASN ASN ASN GLY ASP ASN ALA ASN ALA LYS THR A T A A T A A T G C G A C A A T G C T A A C G C C A A A C A1890 1920		LYS ILE SER ALA ASN GLY HIS ASN VAL THR LEUAAATCTCTGCTAACGGTCACAATGTGACACTA1950 1960 1980		ASN SER ASN THR GLU GLY ASN SER ASP ASN ASN ACAGTAACACTGAAGGTAATAGTGACAATAAC 2010 2020 2040		
GLY A GGCG 1900	: : : : : : : : : : : : : : : : : : :	ALA A GCTA. 1960		THR GACTG,	:	: :
ASN A A T	SER T C A	SER T C T	GLY GGA	ASN A A C	ASIN A A T	•
ASN AAT	LYS ASP A A G A T T	ILE A T C	SER ASP GLY TCTGATGGAA	SER A G T	ALA LYS ASN	• • •
AT A1890	LYS ' A A A 1940	LYS A A A	SER 1 T C T 2000	A C A	۰	2060
• • •	VAL G T T	• • •	THR ACA	: : :	ASP G A T	;
	ASN A A T		GLU GAA		THR ILE) :
	PHE ASN ASN TTAACAATC 1930		SER LYS VAL GLU GCAAAGTGGAAA 1990		祖 A C T	2050
	PHE [T T] 19		LYS A A A A 19		LEU	20
	VAL THR PHE ASN ASN VAL LYS ASP SER STAACTTTAACAATGTTAAAGATTCAA 1930		ASN SER LYS VAL GLU THR ATAGCAAAGTGGAAACA 1990		ALA GLY LEU THR ILE ASP)))
	VAL 3 T ?		ASN 1 A T A		ALA)

	116/235		
VAL THR VAL ASN ASN ASP ILE THR SER HIS LYSTAACAGTAAACAACGATATCACTCTCACAAA2070 2080 2100	ALA G C A 2160	SER T C T 2220	SER T C C
HIS CACA	ASN A A T	LYS A A A A	GLU 3 A A
SER T C T C	THR ILE ASN CCATTAATG	ILE A T T .	LYS ILE GLU AAATTGAA1
THR 3 A C T T 2090	THR A A C C 2150	GLY F G G A 22210	Y LYS TAAA <i>i</i> 2270
ILE A T C	THR A CAA 21.	GLY 3 G T	GLY 3 G T 7
ASP GAT	ASP G A T A	GIN C A A (ASN A A C C
ASN ASN ASP ILE AACBATATCA 2080 20	A SER GLU ARG G T C A G A A A G G A 2120ILE ASP THR LYS ALA ASP THR THR ILE ASN ALAT T G A T A C T A A A G C T G A T A C C A T T A A T G C A T 130 2140 2150	S LEU THR ALA ACTAACAGCTG 2180 VAL THR SER ASP ILE GIN GLY GLY ILE LYS SERTAACAAGTGATATCCAAGGTGGAATTAAATCTTAACAAGTGATATCCAAGGTGGAATTAAATCT2190	TATCACAACCA TATCACAACCA 2240SER THR GLY SER ILE ASN GLY LYS ILE GLU SERGCACAGGTAGCATTAACGGTAAAATTGAATCC2250
ASN P A A C A 2080	 A LYS A A A G 2140	 G ASP I G A T A	 A SER 1 A G C A 2260
VAL THR VAL TAACAGTAA 2070	THR VAL ASN ILE THR ALA SER GLU ARG ACAGTAAATATCACTGCGTCAGAAAGGA 2110ILE ASP THR LYSTGATACTAA	S LEU THR ALA A C T A A C A G C T G 2180 VAL THR SER ASP T A A C A A G T G A '	ILE THR THR TATCACAACCA 2240SER THR GLY SERGCACAGGTAG
THR ACA	GLU GAA ASP GAT	THR ACA THR ACA	THR ACA THR ACA
VAL T A .	A SER 2120 2.1LET T G	S LEU A A C T A A 2180VAL T A A2190	TATC 2240 SER G C
: : :	ALA G C G 	LYS A A A A 2 	ASN A A T 2 2
	THR ACT	VAL G T G	VAL G T A
	ILE A T C 10	ASN A A C	ASP G A T O
	ASN 1 A A T A 2110	GLY A G G C A 2170	GLY <i>P</i> G G G G G G G G G G G G G G G G G G G
	THR VAL ASN ILE THR ALA SER GLU CAGTAAATATCACTGCGTCAGAA 2110ILE ASPTGATA	THR THR GLY ASN VAL LYS LEU THR ALA A C C A C C G C A A C G T G A A C T A A C A G C T 2180 2170 VAL THR SERT A A C A A G T	ASN SER GLY ASP VAL ASN ILE THR THR A A T T C T G G T G A T G T A A A T A T C A C A A C C 2230 2240 SER THR GLY G C A C G G T
	THR ACA	THR ACC	ASN A A T

開

国

盟

VAL

Ŕ

GLY

贸

LYS

G T T 2460

GAG

GGT

ATTGGC

GLU GAG

CAATCAGGC

... C A A G T

...2430

ILE

GLY

GIN

...S**强**

	11//235	
AAGTCTGGCTCTGTAACACTTACCGCAA 2290THR GLU LYS THR LEU THR VAL GLY ASN VAL SERCCGAAAAAACTCTTACTGTAGGCAATGTTCG 2340	GLY ASN THR VAL THR ALA ASN GGCAACACCGTTACTGTACTGCAAATA 2350ARG GLY ALA LEU THR THR LEU ALA GLY SER THRGAGGTGCATTAACCACTTTGGCAGGCTCTACG	ILE ASN GLY THR ASN GLY VAL THR THR A T T A A C G G G A C T A A C T A C C T 2410 2420
	SUBSTITUTE SHEET (RULE 26)

THR GLY LYS THR VAL SER VAL THR ALA .. ACTGGTAAGACAGTAAGTGTTACAGCAA..

<u>-</u>	>	=
	Y	_
<u> </u>	Ξ	<u>)</u>
L	İ	_

	118/235		
ALA G C A 2520	THR A C C 2580	GLY G G T 2640	THR A C T 2700
G G C	THRACC	ASP G A T	LEU
G G T	LYS LEU AAATTA?	GIN C A G (THR ACT (
LYS F A A A 2510	LYS 2 A A A A 2570	ALA A G C T C 2630	GLY A G G C 7 2690
VAL GTT 2	GLY GGCA 25	SER TCAG	THR ACAG
THR ALA GLY SER LEU THR VAL LYS GLY GLY ALA. CTGCCGGCAGCTTAACTGTTAAAGGTGGCGCA. 2490 2520	GLY THR ALA GGAACTGCAA 1540 THR LEU THR ALA SER GLY LYS LEU THR THR CCTTAACTGCATCATCGGGCAAATTAACCACC 2550 2550 2550	THR SER ALA ACTTCAGCCA S600 LYS GLY GIN VAL ASP LEU SER ALA GIN ASP GLY AAGGTCAGGTAGACCTTTCAGCTCAGGATGGT 2610 2620	SER ALA ALA AGTGCAGCTA 1660 ASN VAL THR LEU ASN THR THR GLY THR LEU THRATGTAACAGGCACTCTAACT 2670 2670 2680 2700
LEU TTA 00	SER TCAT	ASP GACC	ASN A A T A 30
SER I AGCT 2500	 ALA S GCAT 2560	ALA 3 C C A CIN VAL P 2 A G G T A G	 LEU ? C T G A 2680
ALA GLY	ALA GCAP THR ACTG	ALA G C C GIN C A G	ALA GCT THR ACA
ALA G C C	GLY THR GGAACTC 540 .THR LEU .CCTTAAA.	THR SER A C T T C A G 600 .LYS GLY . A A G G T C	ALA GCAG VAL GTAA
C T (
• • •	GLU GAA	ILE A T C	ILE ATT
	THR ACA	ASN A A C	GLN C A A
	ALA G C G O	ALA SER SER ASN CTAGCTCAAAC 2590	GLY G G A O
	ASN A T G 2530	SER S A G C T 2590	AIA (G C A G 2650
	LYS ILE ASN ALA THR GLU A A A T T A A T G C G A C A G A 2 2530	ALA 3 C T	SER ILE ALA GLY GIN ILE. GCATTGCAGGACAAAT 2650
	LYS ILE ASN ALA THR GLU A A A A T T A A T G C G A C A G A A 2530	GLU ALA SER SER ASN ILE GAGGCTAGCTCAAACATC 2590	SER ILE ALA GLY GIN ILE A GCATTGCAGGACAAATT 2650

	_)
C	4)
	`	J
,	r	÷
	Ŀ	J
ī	ī	_
	-	_

	120/235		
GLY G G C 2940	GLU G A A 3000	ALA G C T 3060	THR A C A 3120
ARG A G A	ILE A T T	LEU T T A	ASN A A C
ARG LEU GCTTA 30	VAL GTA?	THR ACAT	VAL
ARG 3 C G C 2930	GLU A G A A G 2990	ARG GLU GAGAAA 3050	THR PACG 0
VAL G T G (VAL GLU STAGAAC 29	ARG A G A	ILE ATTA 31
THR ACT G	VAL G T A	GLU GAA	THR ACCA
ARG ASN GAAACA 2920	SER AGTG	GLU GAAG	ASN A A T A
GLU ASN GLY ARG ASN THR VAL ARG LEU ARG GLY A A A A T G G T A G A A C A C T G T G C G C T T A A G A G G C2910 2920 2940	A TYR ILE GIN A TATATCCAGC 2960PRO GLY VAL ALA SER VAL GLU GLU VAL ILE GLUCAGGTGTAGCAAGTGTAGAAGAAGTAATTGAACAGGTGTAGCAAGTGTAGAAGAAGTAATTGAA2970	GAAAGTGAAAG 3020ASP LEU SER ASP GLU GLU ARG GLU THR LEU ALAATTTATCTGATGAAGAAACATTAGCT3030 3040 3050	TGTACGTTTTA 3080 ILE GLU PRO ASN THR ILE THR VAL ASN THRTGAACCAAATAATACCATTACGGTTAACACA3090 3100 3120
GLY GGT	GIN C A G (LYS A A A G SER ASP T C T G A 3	PHE TTT PRO CCA
ASN A A T	ILE ATC GLY GGT	LYS VAL A A A G T G)20 ASP LEU A T T T A 3030	VAL ARG TACGT 80 ILE GLU TGAA 3090
GLU ASN GLY A A A A T G G T2910	S TYR ILE A T A T A T C 2960 PRO GLY C A G G T	J LYS VAL G A A A G T G 3020 ASP LEU A T T T A	A VAL ARG T G T A C G T 3080ILE GLU T T G A A 0
	LYS A A A		ALA G C T
	VAL G T G A	LEU CTT	VAL SER STAAGT(
	ILE GLU TTGAG(2950	G VAL CGTCC 3010	.Y VAL 3 T G T A 3070
	ILE . A T T 29	ARG C G C 30	GLY G G T 30'
	LYS GLU ILE GLU VAL LYS TYR ILE GIN A A G G A A T T G A G T G A A T A T A T	ALA LYS ARG VAL LEU GLU G C G A A A C G C G T C C T T G A 3010	LYS LEU GLY VAL SER ALA A A A C T T G G T G T A A G T G C' 3070
	LYS A A G	ALA G C G	LYS A A A

FIG.23P

			CYS	T G T	3180				* *	T A G	3240
			ALA	$G \subset G$					GLN	CAG	
			LYS	AAG					GLN	CAG	
			GLY	GGT	170				GLY	G G A	3230
			CIN	GAA	3170				ASP	GAT	m
			SE	TCT					ASP	GAC	
			SER GLN VAL THR ILE SER GLU GLY LYS ALA CYS	GTCAAGTGACAATTTCTGAAGGTAAGGCGTGT	0				CYS THR ASN VAL ALA ASP ASP GLY GLN GIN ***	GTACCAATGTTGCTGACGATGGACAGTAG	0
:	A	:	訊	ACA	3160	:	::	:	VAL	GTT	3220
S S	CCAGACCATCAA		VAL	GTG		VAL	G T A		ASIN	AAT	
₹ ?	$C \subset A$		GIN	CAA		ALA	$G \subset A$		選	ACC	
AKG	A G A	3140	Ŗ.	G.	3150	GLY ALA ALA VAL	$G \subset A$	3200	.CYS	G.	3210
芸	ACC		:	:	•	GLY	0 9 b	(-)	:	:	:
其	ACA					ASIN	AAT				
봎	T T T	30				GLY	G G T	96			
	GAG	3130				SE	AGT	3190			
GIN ASN GIU PHE 11HR 11HR ARG PRO SEK	CAAAATGAGTTTACAA					PHE SER SER GLY ASN	TTCTCAAGTGGTAATGGCGCAGCAGTAT				
	CAA					出	TTC				

FIG.24A

PMH1 hmw1A sequence

	122/235		
ASN	A A T 60		ILE A T Ţ
CCCGGATAATG 20VAL ASN ILE VAL LYS GLY THR GLU LEU GIN ASN	TCAATATTGTTAAAGGAACCGAATTACAGAAT 30 50 60		GLU LYS LYS ASN ALA PRO THR LYS THR THR ILE
TEN	T T A		THR ACT
CIU	G A A 50		LYS C A A G
開	ACC		THR ACC
GLY	G G A		PRO C C T
LYS	AAA 40		ALA G C C
G	G T T	: :: :::	ASN A
ASN A A T ILE	ATT	ILE A T T	LYS A A G
PRO ASP ASN C G G A T A A C 20 VAL ASN ILE	AAT	ASP SER ILE 3 A T A G T A T C	LYS A A A
CCCGGATAATG 20VAL ASN ILE VAL	T C	CGATAGTATTG	GLU A G
	• •		•
LEU TTA		ARG AGG	
LEU ; T T G 10		VAL GTT	•
LYS GLU TRP LEU LEU ASPAAAAGAGTGGTTGTTAGA(ASP LEU VAL VAL ARG GLYGATTGTTAGGGGG	
GLU AGAG		LEU 'TTG	
LYS A A A		ASP G A T	

380

370

ASN LYS VAL ASN VAL THR THR ASP ILE A A T A A A G T A A T G T T A C T A C A G A T A T T A 190 ASN VAL TYR ASN GLY ALA LEU THR LEU HIS SER 200 AT G T T T A T A A T G G A G C A T T A A C G T T A C A C T C A	LU LYS ASN GLY ASN LEU 1 A A A A A A T G G T A A T T T A A 290	ILE LYS ALA GLY SER TRP VAL ASP VAL ATTAAAGCAGGTAGCTGGGTTGATGTTC 310HIS LYS ASN ILE THR LEU GLY GLU GLY PHE LEUATAAAATATCACACTTGGCGAGGGTTTTTTGATAAAAATATCACACTTGGCGAGGGTTTTTTG	ASN ILE THR SER GLY ASP ILE ALA PHE A A T A T T A C T G G T G A T A T C G C C T T C G							
A A	SUBSTITUTE SHEET (RULE 26)									

A A G 600

AAAAACO

AAA

CAAACTACG 580

... CGATTAAT

GLN

TCCGGAGTAGTAA... 560

VAL

SE

ASIN

ILE

ALA

ILE A T T

3 A T T A A C A 550

909099

 $G \subset T$

ASIN

124/	235
------	-----

				12	4/	233)						
GIEN	CAA	420			ASIN	AAT	480				SER	AGT	540
ALA	C L				LEU	TTA					ILE	ATT	
当	ACC				SER	$T \subset T$					ASN	AAC	
IIE	АТТ	410			VAL	GTA	470				HIS	CAC	530
置	ACC				ASIN	AAT					景	ACA	
国	C T A				ASN	AAT					盟	LTT	
GLU LYS GLY ASN ASN LEU THR ILE THR ALA GLN	A A A A A G G T A A T A A T C T A A C C A T T A C C G C T C A A	400			LYS GIN LEU ARG LEU ASN ASN VAL SER LEU ASN	A A C A A C T T A G A C T T A A T A A T G T A T C T T T A A A T	460				ALA ASN GIN ASN ASN PHE THR HIS ASN ILLE	CAAATCAAATAATTTACACACAATTAGT	520
ASIN	AAT	40	: ,	A	ARG	A G A	46	:	.: G	:	ASIN	AAT	52
GLY	G G T		GLY	G G A	PEI IEI	C T T		HE.	ATT		GLN	CAA	
LYS	AAA		ASP	G A C	GIN	CAA		盟	TTT		ASIN	AAT	
GIO.	A A	390	SXT	A A A 440	.LYS	. A A	450	ASN	AAC	200	.ALA	. CA	510
•	•	•	SER ASN LYS ASP GLY	A A T	:	•	:	PEN TEN	ΤΤΑ		:	:	:
			段 。	T C T				GLY	G G T				
			音,	A A C C 430				ALA	GCA	490			
			田田	A T A L				CLY	G G T	Ţ			
			GLY ASN ILE THR	G G A A T A T A A C C T C T A A T A A A G A C G G A A 430				GLY THR GLY ALA GLY LEU ASN PHE ILE	GGAACAGGTGCAGGTTTAAACTTATTG				
			GLY	d G G				GLY	G G A				

125/235

TGACTCTTACT GACTCTTACT G20TRP ASN VAL SER THR LEU THR LEU SER ASN ASPGGAACGTATCTACTTTACTTTAGCAATGAT E030 E120 E120 E120 E120 E120 E120 E120 E12
--

L	I,	1
7	J	1
C	1	J
	·	•
(r	5
_	=	_
L	ı	_
_		-

	126/235		
PRO C C A 840	ASP G A T 900	SER C T 960	ASN 1 A T 1020
ARG A G A (PHE I T C (VAL 3 T T 1	II.E
ASN A A C i	PHE	ASN A A C (GLU 3 A A P
PRO C C T 830	VAL G T G T 890	ILE ATT? 950	PHE [T T T C 1010
THR ACT	SER T C T	SER	ALA G C T 1
LYS PRO ASN GLU LYS THR THR PRO ASN ARG A A C C A A A T G A G A C A C T C C T A A C A G A 810 820	SER ASN ILE TCTAATATTT 860 SER VAL THR GLY GLY SER VAL PHE PHE .CGGTCACTGGCGGAGGTTCTGTGTTTTTTC. 870 880	GLY LYS GLY GGTAAAGGGA 920 THR GLU LEU LYS MET ASP SER ILE ASN VAL SER CTGAGCTAAAGATGGATTCAATTAACGTTTCT 930 940	LEU ASN SER TTAAATTCCC 980HIS VAL ARG LYS TYR ASN ALA PHE GLU ILE ASNATGTTCGCAAGTATAATGCTTTTGAAATCAATATGTTCGCAAGTATAATGCTTTTGAAATCAAT990 1000
LYS A A G A 20	. GLY C G G A (880	S MET G A T G (940	TYR T A T A
ASN GLU AATGAGA 821	 G.C.G 880	 LYS A A G A	 LYS T A A G T 1000
ASN A A T	ILE ATTT THR ACTG	GLY GGGB LEU CTAA	SER TCC ARG CGC
PRO C C A A	ASN AATA VAL GTCA	LYS A A A GLU G A G	ASN AAT' VAL 3TT(
LYS . A A C	SER ASN ILE TCTAATATTT 860 SER VAL THR GLY . CGGTCACTGG	GLY LYS GLY GGTAAAGGGA 920 THR GLU LEU LYS . CTGAGCTAAA	LEU ASN SER TTAAATTCC 980 .HIS VAL ARG .ATGTTCGC
: : :		TRP T G G 	THR ACC
	PHE T T T T	LEUCTT	LEUCTT
	GIN C A G 0	ASN A A C	ASN A A T O
	PRO ILE GLN CAATTCAGT 850	TYR ALA ASN LEU PACGCTAACCTTT 910	GLY SER ASN LEU THR GCTCTAATCTTACO 970
	PRO C C A	TYR T A C	GLY 3 G C '
	LEU PRO ILE GIN PHE LEU CTACCAATTCAGTTTTTA 850	ILE TYR ALA ASN LEU TRPATACCTTTGG910	SER GLY SER ASN LEU THR AGCGGCTCTAATCTTACC 970

ILE

IEO Е ⊱⊣

	127/235		
SER A G T 1080	GLY G G C 1140	ILE A T C 1200	
GCAACTAATT 1040 SER ASN LEU ARG GLN THR SER ASP SERCAAATTTCAACTCAGACGTCAGATAGTCAAATTTCAACCTCAGACGTCAGATAGT	ASN ASN ALA CAATAATGCCA 1100 ILLE ASN SER THR HIS ASN ILE SER ILE LEU GLYTCAATTCAACCCACACATATCCTTGGGCTCAATTCAACCCACACATATCTTGGGC1110	GLY GIN ASN GGACAAAACT 1160 SER SER SER ILE MET GLY ASN ILE ILE CAAGCAGCATTATGGGGAATATCATCATC 1170 1180 1180	
SER T C A G	ILE A T C	ILE A T C	
THR 3 A C G 1070	SER A T C C 1130	ASN 3 A A T 1190	
GLN CAG	ILE ATA 1	GLY GGGG	
ARG A G A	ASN A A C	MET A T G	
LEU CTC	HIS CAC	ILE ATT 30	
GCAACTAATT 1040 SER ASN PHE ASN LEU ARG GLN THRCAAATTTCAACCTCAGACGT1050 1060	 A THR H A C C C 1120	ASN A A C T SER SER ILE MET GLY ASN ILE A G C A G C A T T A T G G G G A T A T C A 1180	: : :
CGCAACTAATT 1040SER ASN PHE ASNCAAATTCAA	ALA GCC SER TCA	GLY GLN ASN 5 G A C A A A A C 160 SER SER SER C A A G C A G C 1170	GLU GAA
ASN ASN A A T	ASN ASN ALA A T A A T G C C 00 ILE ASN SER T C A A T T C A	GIN C A A SER A G C	THR LEU GLU CGCTAGA 20
C G C A 1040 SER C A 1050	C A A T 7 1100 ILE T C 7 1110	C G G A C A A A 1160SER SER C A A G C A C A A G C A	THR T A C G 1220
	ARG C G C	GLY GGC	VAL GTT
ATAT	TYR 3 T A C C	LEU FCTC	ASN AAA
AAAGACTTAACTATAAA 1030	PHE ARG ASN GLY TYR ARG ASN ASN ALA T T T C G T A C C G C A A T A A T G C C A 1090ILE ASN SER THRT C A A T C A A C A	GLY ASN VAL THR LEU GLY GLY GIN ASN GGCAACGTCACTCTCGGCGGACAAACT 1150 1160SER SER SER SER SER 1CAAGCAG	LYS ARG ALA ALA ASN VAL THR LEU GLU A A G C G A G C A A T G T T A C G C T A G A G 1210
10 T T 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ASN I A A (VAL CGT(ALA A G C 7
A G A (ARG I C G .	ASN	ARG 3 C G 7
AAA	PHE TTT	C C (C	LYS A A (

C	r	5
3	ノ	J
(_	5
L	Ī	_

	128/235		
LYS A A G 1260	GLU G A A 1320	THR A C C 1380	II.E A T A 1440
ALA ASP ASN SER HIS ASN SER ASP ASN VAL LYSCCGATAATAGTCACAATTCTGACAACGTAAAG	GLY ASN LEU FGCAACTTGA 1280 THR VAL GLU GLY ASN LEU SER LEU ILE GLY GLU CCGTTGAGGGAATTTAAGTTTAATTGCGAA 1320	ASN LEU SER 1340ILE GLU LYS GLU ALA ILE PHE LYS GLY LYS THRTGAAAAGCCATCTTTAAAGGAAAACCTTGAAAAGCCATCTTTAAAGGAAAAACC13501350	THR GLY ASN LACCGGCAACT 1400 THR ASN ASN GLY THR ALA GLU ILE ASN ILE TTACCAATAATGCCACTGCCGAAATTAATA 1410 1420 1440
ASN A A C	ILE A T T	GLY G G A	ILE A T T A
T G A C 1250	SER LEU ILE GTTTAATTC 1310	LYS F A A A 1370	GLU G A A 1430
SER TCTG	SER A G T	PHETIT	THR ALA GLU CTGCCGAAA
ASN TAA:	LEU	ILE A T C	THRACT
ASN SER HIS ASN ATAGTCACAATT 1240	GLY ASN LEU FGGCAACTTGA 1280 THR VAL GLU GLY ASN LEU CCGTTGAGGGAATTTA?	ASN LEU SER 1340ILE GLU LYS GLU ALA ILE PHE LYS GLYTGAAAAGCCATCTTAAAGGAA 1350 1350 1360	GLY ' G G C A 20
SER 1 A G T	 A GLY G G G	 GLU P G A A G	T ASN A A T T
ASN ASN	ASN LEU A A C T T C VAL GLU S T T G A G	SER TCC LYS AAA	ASN A A C ASN ASN
ASP G A T	ASN ASN VAL VAL	ASN LEU SER A T C T C T C C 40 ILE GLU LYS T G A A A A A 1350	THR GLY 1 C C G G C 7 100 PHE THR T T A C C A 1410
ALA ASP C C G A T A	GLY T G G C 7 1280 THR C C G	ASN 2 A A T 1340 ILE T T	THR GLY ASN 1A00 PHE THR ASN ASN TTACCAATAATC1410
• • •	LEU CTT	ASN GLY ACGGC	ILE ATC
	ASN A A T C	ASN A A C	LEU ASN ILE TAAACAT
	ILE ATAA 10	ILE ATTA 0	LEU CTA O
	THR I ACTA 1270	ASN I A A T A 1330	SER L A G C C 1390
	ASP ARG THR II.E ASN LEU GLY ASN LEU GATAGAACTATAAATCTTGGCAACTTGA 1270THR VAL GLU GLYCCGTTGAGGG	ASN ALA ASN ILE ASN GLY ASN LEU SER A A T G C A A T A T T A A C G G C A A T C T C C A 1330ILE GLU LYS GLUT T G A A A A G A A 1350	LYS ASP SER AGGACAGCO 1390
	ASP GAT	ASN A A T	LYS ASP SER LEU ASN ILE THR GLY ASN A A G G A C A G C C T A A A C A T C A C G G C A A C T 1390PHE THR ASN ASNT T A C C A A T A A 7

EEI

VAL

VAL.

GLY

GIN

SER GIN AGCCAA(

AGT SEG

	129/23	5		
THR A C T	1500	ASN A A T 1560	GIN C A A 1620	
THR)	GLY G G G	SER T C G (
ILE ATCA) 1	GIN C A A	ILE	
ASN	1490	ASN LYS GLN GLY ACAAACAAGGG 1550	ASN 2 A A T A 1610	
LYS LEU		ASN AAC	GLY G G C A	
TCTTGGCGATA 1460 ILE THR ASN ASP GLY LYS LEU ASN ILE THR THR	G G	1520 ILE ARG GLY ASP ILE ILE ASN LYS GIN GLY ASN TCCGCGAGATATAATTAACAACAAGGAAT 1530 1540 1550	TAATAGTAATG TAATAGTAATG 1580 ALA GLU ILE GLU ILE GLY GLY ASN ILE SER GINCTGAAATTGAAATTGGCGCAATATCTCGCAA1590 1600 1600	
THR ASN ASP GLY) 0 0 0 0 0 0	ILE ATAA 40	 ILE A A A T T 1600	
ASP	1480 1480 	ASP] GATA 1540	 GLU GLU 160	: : :
G A T ASN	AA I ILE AT I	ARG GLY GCGGA	ASN A A T ILE	SER
G G G	SER AGG	ARG C C G C	SER FAGT GLU GLU	SER I T C T
TCTTGGCGATA 1460 ILE THR ASN ASP	1470 1 INS SER ILE	1520 ILE T C	T A A T A G T A A T 1580ALA GLU ILE C T G A A A T T T1590	: ILE SER SER CATTTCTTCTG
_	GIN FCA7		'	~
7 L D D	GLY C G G T C		ASP G G A (LEU TCTC?
3 A G T (SER GAGC(1510	ILE THR TTACG(1570	Y ASN 3 TAATC
A G G 7	AIA LYS S C C A A G	T	I ILE TAT 1	J GLY AGG
AGCCAAGGAGTGGTAAG	HIS ALA LYS SER GLY GIN		LEU ASN ILE THR ASPAST TTAAATTACGGACAA 1570	LYS GLU GLY ASN LEU THI AAAGAAGGTAATCTCAC 1630
A G	HIS C A C (AT T	LYS A A

7	4
C	V
_	
Ĺ	ני
_	_
L	L

	130/235		
AFP LYS VAL ASN ILE THR LYS GIN ILE THR ILE A T A A A G T C A A T T A C C A A A C A G A T A A T C 1650 1660 1680	LYS AIA GLY VAL ASP GLY GLU SER SER A A G C A G G C G T G G G G G G G T T C T A 1690 SER SER THR AIA SER ASP AIA ASN LEU THR OF T C A G C A G C A G T G A T G C C A A T C T A A C C C A G C A G T G A T G C A A T C T A A C C C A G C A G T G A T G C A A T C T A A C C C A G C A G T G A T G C A A T C T A A C C C A G C A A G T G A T G C A A T C T A A C C C A A G T G A	ILE LYS THR LYS GLU LEU THR LEU THR ATTAAAACCAAAGAGTTAACAG 1760 ASP ASN LEU ASN ILE SER GLY PHE ASN LYS ALA ACAATCTAAACATTTCAGGTTTTAATAAAGCA 1800	GLU ILE THR ALA LYS ASP ASN SER ASP GAAATTACAGCTAAAGATAACAGTGATT 1810 LEU ILE GLY LYS ALA SER ASP ASN SER TAATTATTGGCAAGGCTAGCAGTGACAACAGT 1830 1840 1850
	VAL AGTTG	LYS GAAAG50	ALA L G C T A .
	A GLY A G G C 169	LYS THR AAACCA 1750	TACA 181
	LYS ALA AAAGC	ILE LYS ATTAA	GLU ILE 3 A A A T
	·	· -	O

PCT/CA99/00938

AAT 2040

₽

G

Ø

E

Z,

贸

ASN

WAL

黑

WAL

: : AAG ی ASP ď ⊟ 出 CCTT 胃 Ø Ø WAL ⊣ G Ø GEN Ø ں AA LYS Ø ⊟ ALA ت G ⊱⊣ ASIN AA

1870

U L Ø LYS 段

CAAAA ۳ ⊣ G B AAA ...VAL

CAC

⊱

Ø

Ø

 \mathcal{C}

ტ ტ

TCAGCT

ASIN

GLY

段

1900

..1890

GAA T G G

AGCAAA

⊱⊣

K

Z,

Ø

CAC

Ø

Ø

C

⊱

Ø

LYS

段

ASIN

园

WAL

ASSN

131/235

段 ASIN 段 ⊣

置:::

ASP G A ں G Z, ₽ Ø Ø

ပ ပ

GGT

AAC

GGA

AGCACC

r L

G

E

ASIN

GLY

温

贸

Ä ACT TTA 国 TCGGC GLY Ø ₽ ASIN ø Ø

GACAAC

ΑT

G

 \mathcal{O}

A G

ASN

ASP

ASP LYS ALA 段 图:::

CGGTA Ø T A G ⊣ AAGA Ø . ဗ Γ C C

AT ⊢ ASN AA ACAGTA WAL 訊 AAA LYS A C HIS ی $^{\perp}$ 贸 ₽ ပ ပ K

2050

 \mathcal{O}

⊣ Ø

ILE

_	<	
7	4	
C	V	
,	'n	
	י	
ī	ī	
_	_	

	132/235		
ALA G C A 2100	ILE A T T 2160	VAL G T T 2220	ASN A A T 2280
SER ALA SER GLU GLY GLY ILE THR THR LYS ALA CTGCATCAGAGGTATCACTACTAAGCAG2070 2080 2100	THR THR GLY 3 A C C A C A G G T A 2120 SER VAL GLU VAL THR ALA LYS THR GLY ASP ILE G C G T G G A A G T A A C T G C T A A A C A G G C G A T A T T 2130 2140 2150	LYS THR VAL LAAGACAGTAA 2180 SER VAL THR ALA SER THR GLY ASP LEU THR VALGTGTTACAGCAAGCACTGGCGATTTAACTGTT 22002190	SER ALA THR LAGTGCGACAG 2240 GLU GLY ALA ALA THR LEU THR ALA THR GLY ASNAAGGAGCTGCAACCTTAACCGCAACAGGGAATAAGGAGCTGCAACCTTAACCGCAACAGGGAAT2250
THRACT	G G C	LEU TTA	THR ACA
ILE THR THR TCACTACTA 2090	THR A A C A 2150	ASP 2 G A T 2210	ALA 2 G C A 2270
ILE	THR THR GLY 3 A C C A C A G G T A 2120 SER VAL GLU VAL THR ALA LYS THR GLY G C G T G G A A G T A A C T G C T A A A C A G C C 2130 2140	SER THR GLY ASP GCACTGGCGATT	THR LEU THR ALA CCTTAACCGCAA 2270
SER ALA SER GLU GLY GLY C T G C A T C A G A A G G A G G T A 2070 2080	ALA GCT	THR	LEU T T A
LU GLY A A G G A 2080	 AL THR TAACT 2140	 .A SER 2 A A G C 2200	 .A THR : A A C C 2260
GLU AGAA	 T A VAL A G T A	LYS THR VAL LAAGACAGTAA 2180 SER VAL THR ALAGTGTTACAGCAA21902190	A G ALA I G C A 226
SER	GLY GLU GLU	LYS THR VAL A G A C A G T 1 80 SER VAL THR G T G T T A C 1 2190	SER ALA THR GTGCGACA 40 GLU GLY ALA A A GGAGCT 2250
ALA G C A	THR ACAO VAL GTGO	THR 3 A C A VAL G T T	ALA GC G GLY G G A
SER C T	GLY THR THR ILE ASN ALA THR THR GLY GGCACACATTAATGCGACCACAGGTA 2110SER VAL GLU VALGCGTGGAAGT.	SER GLY THR ILE SER GLY LYS THR VAL AGCGGTACGATTCCGGTAAGACAGTAA 2170 2180SER VAL THR ALAGTGTTACAGC	ARG LYS ALA ALA THR ILE SER ALA THR A G G A A A G C T G C A A C C A T T A G T G C G A C A G 2230GLU GLY ALA ALA ALA ALAA A G G A G C T G C
• • •	ALA G C G	SER GIY CCGGI	ILE ATT
	ASN AATC	SER T C C	THR ACC
	ILE A T T A 10	ILE ATTT 70	ALA G C A 30
	THR THR 1 CAACCA 2110	THR A C G 21'	ALA P G C T G 2230
	THR ACA	SER GLY THR GCGGTACGA 2170	ARG LYS ALA ALA THR GGAAAGCTGCAACC 2230
	G G C	SER A G C	ARG A G G

, G C T A T T 2460

GCAACCAGTGGCACC 2440 2450

ALA

E

開

GLY

贸

開

ALA

LYS

田田

... ASIN

A A G (

... ACATT

GGT

ASP

国 日

LYS

ALA

ASP

LYS

ALA

G C T A A G C T A G A T 2480

AAAGACO 2470

ASN A A C

3.24L

			ILE THR SER THR LYS GLY GLN VAL ASP LEU SER	TCACTTCAACTAAGGGTCAGGTAGACCTTTCA	2340				ILE SER ALA ALA ASN VAL THR LEU ASN THR THR	TTAGTGCAGCTAATGTGACATTAAATACCACA	2400				
			ASP	GAC					ASN	AATA					
			VAL	G T A	2330				EEG	$\mathbf{T} \ \mathbf{T} \ \mathbf{A}$	2390				
			GIN	CAG					强	$A \subset A$					
			CLY	G G T					VAL	G T G					
			LYS	AAG	30				ASIN	AAT	20				
:	A	:	国	ACT	2320	•	A	:	ALA	$G \subset T$	2380		•	A	:
S E S	AGC		SER	T C A		GIN	CAA		ALA	GCA			SER	T C A	
SE	TCT		强	ACT		Y, T.)	GGA		Ŕ	AGT			GLY	GGT	
ALA GLY SER SER	GCCGGTTCTAGCA	2300	· IIE	T C	2310	TIFE ALA CILY GIN	GCTCAGGATGGTAGCATTGCAGGACAAA	2360	TITE	. T T	2370		VAL GLU GLY SER	GGCACCTTAACTACTGTAGAAGGTTCAA	2420
ALA)) b		•	•	•	<u>Н</u>	ATT		•	•	•		VAL	G T A	
GIN	ACCTTGACTACTGAA					S. T.	AGC						强	ACT	
訊	ACT	2290				۲۱:	GGT	2350					開	ACT	2410
置	BACT	22				ASP	GAT	23					TEG	TTA	24
FE	TIG					ALA CIN ASP	CAG						强	ACC	
置	ACC					ΔI.Δ	G C 1						GLY	0 9 9	
					~		-1	~ ~	CU	===	(D)	=	26	`	

133/235

FIG.24N

			-
	134/235		
THR ALA SER GLY ASN ARG THR GLU VAL ASN ALA . CGGCATCAGGTAACCGTACAGAAGTAAATGCA .2490 2500 2500	GLY SER VAL GGTAGCGTGA 1540 THR ALA LYS THR SER SER ASN VAL ASN ILE THR. CTGCGAAAACCTCAAGTAATGTGAATATCACC 2550	ASN GLY LEU A A T G G G T T A A 1600 ASN ILE ILE SER GLU ASN GLY ARG ASN THR VAL . A T A T C A T T T C G G A A A T G G T A G A A A C A C T G T G . 2610 2620 2640	GLU G A A 2700
ASN A A T	ILE A T C	THRACT	VAL G T A
VAL G T A A	ASN A A T P	ASN A A C A	SER AGCG
THR GLU CAGAAC 2510	VAL r G T G 2570	ARG A G A A 2630	ALA 1 G C A A 2690
THR ACA	ASN A A T C	GLY 3 G T A	VAL 3 T A G 26
ARG C G T A	SER AGTA	ASN A A T C	GLY 3 G T (
ASN A A C (SER ICAA	GLU 3 A A A	PRO C C A G
GLY 7 3 G T A 2500	 A THR S A.C.C.T 2560	A SER C T C G G 2620	 A GIN I C A A C 2680
SER TCA(VAL GTG LYS AAA	LEU TTA ILE ATT	VAL GTG ILE
ALA G C A T	GGTAGCGTGA S540 THR ALA LYS THR . CTGCGAAAAC	ASN GLY LEU A A T G G G T T A A 2600ASN ILE ILE SER A T A T C A T T T C 2	ILE ASP VAL A T T G A T G T G A 2660 LYS TYR ILE GIN PRO GLY VAL ALA SER VAL GLU A A T A T A T C C A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G A A C C A G G T G T A G A A C C A G G T G T A G C A A G C G T A G C A A G C G T A G A A C C A G G T G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C G T A G C A A G C C A G G T A G C A A G C G T A G C A A G C C A G G T G T A G C A A G C G T A G C A A G C C A G G T A G C A A G C C A G G T G T A G C A A G C G T A G C A G C T A G
. C G (GLY G G T A 2540 THR C T G	ASN A A T C 2600 ASN A T A	ILE A T T C 2660 LYS A A T
: : :	<u> </u>		
	GLY G G T	THR A C A	ARG GLY LYS GLUAGAGGAA
	SER A G T G O	SER 'A G C A O	GLY G G C O
	AIA 8 G C A A 2530	LEU S TTAA 2590	ARG (A G A G A G A G A G A G A G A G A G A
	THR ASN ALA SER GLY SER CTAACGCAAGTGGTTCT 2530	GGGATTTAAGCACAATA 2590	ARG LEU ARG GLY LYS GLU GCTTAAGAGGCAAGGAA 2650
	THR CTA	GLY : G G G	ARG GCT

PCT/CA99/00938

FIG.24N

GLU ALA LYS ARG VAL LEU GAAGCGAAACGCGTCCTTG...

GLU VAL ILE GAGGTAATT

	135/235		
ARG A G A 2760	11.E A T T 2820	GLU G A A 2880	
GLU GAA	PRO ASN ASN ALA CCAAATAATGCCA 2810	SER T C T (
GLU GAA	ASN A A T	ILE A T T	
ASP F G A C 2750	ASN A A A T 2810	11.E 3 A T A 2870	
SER TCT 2	PRO C C A	VAL GTG	
2720 GLU LYS VAL LYS ASP LEU SER ASP GLU GLU ARG AGAAAGATTTATCTGACGAAGAAGA 2730	GLY VAL SER IGGTGTAAGTG 2780ALA VAL ARG PHE VAL GLU PRO ASN ASN ALA ILECTGTACGTTGAGCCAAATAATGCCATTCTGTACGTTGGAGCCAAATAATGCCATT2790	GLU PHE THR I.G.A.A.T.T.A.C.A.A 2840THR ARG PRO SER GIN VAL ILE ILE SER GLUC.C.A.G.A.C.G.T.C.A.A.G.T.C.A.A.T.T.C.T.G.A.A. 2850 2850 2850	
ASP GAT	 HE VAL T C G T T 2800	SER A G T	
LYS A A A A A B 2740	 G PHE 丁丁C 28(A SER SER 1 T C A A 2860	: : :
VAL G T A	GLU THR LEU ALA LYS LEU GLY VAL SER GAAACTAGCCAAACTTGGTGTAAGTG 2770AIA VAL ARG PHECTGTACGTTT	THR ILE ASN THR GIN ASN GLU PHE THR ACGATTACACAAAATGAATTACAA 2830THR ARG PRO SERCCAGACGTC	GLY LYS ALA CYS PHE SER SER GLY ASN GGTAAGGCGTGTTTCTCAAGTGGTAATG 2890 2890
LYS A A A	VAL GTA VAL GTA	PHE TTT ARG AGA	G G T
2720 GLU A G 7 2730	T G G T 2780 ALA C T 0	GLU F G A A 2840 THR C C	SER A A G T 2900
	CTT	ASN A A T	SER T C A
	LYS A A A	GIN C A A	PHE TTCT
2710	GLU THR LEU ALA LYS LEU AAACACTAGCCAAACT 2770	THR ILE ASN THR GIN CGATTAATACACAA 2830	GLY LYS ALA CYS GTAAGGCGTGTT 2890
27	LEU CTA 27	ASN A A T 28	ALA G C G 28
	THR A C A	ILE A T T	LYS A A G
	GLU GAA	THR A C G	GLY GGT

136/235

ASIN GTATGT ME ... G C G C A G C A (

> GLY GLN PRO *** GGACAGCCGTAG 2950

A A C 180

G C A 1 170

ASN A A C (

GAG

VAL

THR

...TE

国

LYS

AAGCCGACA

GAC

T C C A (

CA

ACAAAC

TGTT(160

ACA.

... C A ... 150

AA

E

<	1	_
L	C)
C	1	J
(_	έ.
<u>ر</u>	_	2
Ĺ	l	_

PMH1 hmw2A sequence

P PRO ASP ASP C C G G A T G A T G 20 VAL THR ILE ALA ALA GLY ALA PRO GLY ARG ASN TAACTATTGCCGCAGGCGCCCAGGACGTAAC 30 40 60	LYS A A A 120
CCGGATGATG CCCGGATGATG 20VAL THR ILE ALA ALA GLY ALA PRO GLY ARG ASNTAACTATTGCCGCAGGCGCCCAGGACGTAA (6) 30 40 60	CTTTTTCCCA 80 THR GLY ARG GLY ASP ALA SER ASN ALA LYSCTGGAAGAGGGATGATGCTAGTAATGCAAAACTGGAAGAGGGATGATGTAGTAATGCAAAA90
GLY	ASN
GGA	A A T
PRO	SER
C C A	AGT
50	110
ALA	ALA
G C G	G C T
G G C	ASP G A T
ALA G C A 10	7 ASP GGAT 100
G	A
ALA	G. G. G.
G C C	10
ASP PRO ASP ASP A C C G G A T G A T G 20VAL THR ILE ALATAACTATTGC 30	ASP PHE PHE PRO ACTTTTTTCCCA 80 THR GLY ARG GLY CTGGAAGAGG
ASP	PHE
GAT	TTT
THR	GLY
ACT	GGA
ASP PRO ASP ASP ACCGGATGA 20VAL THR ILETAACTATT 30	ASP PHE PHE PRO A C T T T T T C C (80 S) THR GLY ARG C T G G A A G A 90
St A	ASF A
LEU	ASP
TTA	G A C
LEU	VAL
TTG	GTA
10	70
TRP	SER
TGG	TCA
LYS GLU TRP LEU LEU A	ASP GLY SER VAL ASP 7
A A A G A G T G G T T G T T A G	GATGGTTCAGTAGACG
10	70
LYS	ASP
A A A	GAT
SUBSTITUTI	E SHEET (RULE 26)

出

ASIN

出

ASIN

138/235

T T 360 300 G Ø Ø G ₽ G ASIN T G Ø Ø Ø Ø 060 \mathcal{C} IE Ø AGC AGC GAC 350 TCC C ⊣ 段 贸 G G Ø Ø TAC⊱ ASIN æ Þ ⊱ Ø E WAL Ø \mathfrak{O} Ø Ø 280 340 E Ø \vdash 開 Ø K Ø \mathcal{O} \mathcal{O} K E AAAAA ⊣ Ø ⊣ C T C ACCACT E ø C [H] ATAA GAAA ACC TAACC . G C G 210 選 A C TCT TA ⊱ 段 HE E E K Ø CAC E \mathcal{O} ASIN Ø Ø Ø Ø ပ ပ ⊱ Ø ALA HE Ø <u>ပ</u> G ø K 190 ⊱ E→ ⊱ ASIN ₽ \mathfrak{O} Z, Ø ⊱ G Ø ى ပ ပ ⊱ ⊱ G A Ø Ø Ø

(•)
Ļ	?	į
(•	i
<u>(</u>		5
Ĺ	ĺ	_

139/235		
LYS A A A 480	LEU C T A 540	TRP T G G 600
ILE A T C	SER T C C	SER T C T
VAL G T A	VAL G T C	ASN A A C T
SER T C T 470	ASN A A C 5	SER AGCA 590
GLU 3 A A	ASN A T A	SER AGTA
SER ICA(PHE I T T I	ILE A T C A
VAL G T A C	ARG A G G C	ASN A A T A
 GIN 3 A A (3 PHE F T T 7	A LEU ARG PHE ACTGCGCTTCG 560 VAL GLY GIN LYS ASN ILE SER SER ASN SER TRP TTGGTCAGAAAATATCAGTAGCAACTCTTGG 570 580 600
GLU G A G Z ARG C G C (VAL GTA (GLY	PHE TTC(GLN GLN
ASN A A T GLY G G G	GLY G G T GLU GLU	A C T G C G C T T C 560 VAL GLY GLN T T G G T C A G
GLY G G A 440 .LYS . A A .	SER T C A 500 .GLY .GG G	JY LEU 3 A C T G (560VAL T T G
ALA G C C 	THR ACC 	GLY G G A
LYS A A A A	ILE A T C	ALA G C A O
GLU G A G O	VAL G T T O	G G C C
PHE TTC 43	GLY G G A 49	VAL (GTTG
ALA G C C	GLN C A G	ASN GLY VAL GLY ALA GLY A A T G G C G T T G G C G C A G G A 550
VAL G T A	ALA G C C	ASN A A T
	GIN VAL SER GLU SER VAL ILE LYS CAAGTATCAGAATCTGTAATCAAA 460 470 480	C G A A G G A A C C A A C C A A

140/235

ARG GLU ASN THR ILE LYS ASN ARG PHE A G A G A A A C A G A T T C G 620 ASP GLY ASN LEU ASN ILE SER GLY LYS VAL ASNAT G G G A A T T T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T C T C A G G A A A G G T A A A T A T G G G A A A G G T A A A T A T	VAL SER MET ASP VAL SER GLY THR LYS G T T T C A A T G T A T C C G G G A C A A A G T 670TRP HIS THR ARG ILE ASN GLY ARG THR TYR TRPG C A T A C A G A A T T A A C G G C C C C T A C T G G 720	ASN VAL THR THR LEU ASN VAL ALA SER A A T G T A A C G T T G C C T C A G 740 GLY SER SER PHE ASN LEU SER ILE ASP ALA SER G T A G T A G T T T C A A T C T C A G T A T C G A C G C C A G T 780	GLY ILE SER GLY ASN GLN ASP ASP GGAATTTCTTCAGGTAACCAGGACGACA 790 800
U THR CAC(610	G G A 7	C A C 730	C SER TTC1 790
A A A A	A A T	A A C	SER TTC
GLU GAAA	SER T C A A	VAL G T i	ILE A T T T
A G A G	VAL GTTT	ASN A A T G	GLY GGAA

PCT/CA99/00938

ACA

ILE T A

> G G C 830

ASN A A T

TTA

GLY

ARG

ASIN

...ILE

FIG.25E

AGGGGT: 820

AAT.

THR ACA

... T A i

ASIN

ASIN

GEU

GLY

141/235

VAL G T T 960	SER
ASN ILE SER VAL A A T A T T T C A G T T 950	HIS THR THR SER
ILE A T T	THR
ASN A A T 950	HIS

			_
			HIS
			出
			SER
			SER
:	: ز	:	SER
CIN	G A A		ALA
阳	I I I		ASN
ASIN	AACTGTCAACTTTGAA	980	LEU
VAL	GTC		•
開	$A \subset T$		
GLY	GAGGA	970	
GLY	G G A	9	
GLY	G G A		
LEU	ТТА		

930

TAGCACCCACACAACTTCT 1010 CATC 1000 H ည ပ

ASN A A

T C A A A T T T T A T C A A A A T T T T A 1040 ASN VAL SER GLY GLY SER LYS LEU ASN LEU LYS A T G T C T C A G G T G G G T C A A A T T A A A T C T C A A G 1050 1060 1080	TACCGCTTTTT 1100 LEU ILE LYS ASN ASN LEU THR LEU ASN ALA THRTAATAAATAATTTAACTTTAAACGCTACTTAATAAAAATAATTTAACTTTAAAGGCTACT1110	TAAACAGGTTG 1160GLU GLY THR ASP SER ARG ILE GIN LYS GLY VALAGGGTACCGATTCGCGCATTCAAAAGGTGTTAGGGTACCGATTCGCGCATTCAAAAGGTGTT1170	
SER LYS LEU CAAAATTAA 1070	THR ALA PHE TACCGCTTTTT 1100 LEU ILE LYS ASN ASN LEU THR LEU ASN TAATAAATAATTAACTTTAAACG TAATAAAAAATAATTTAACTTTAAACG	S ILE GIA CATTCA 1190	
 JLY GLY SE G T G G T C 1060	 .SN ASN LE ATAATTT 1120	 .SP SER AR .ATTCGCG 1180	
TTTTA SER GL	PHE T T T T LYS AST A A A A A A A A	VAL 3 G T T G THR ASP	GLU FGAAG
SER GIN ASN PHE T C T C A A A A T T T T A 1040ASN VAL SER GLY A T G T C T C A G G T C 1050	ASN THR ALA PHE A A T A C C G C T T T T T 1100 LEU ILE LYS ASN T A A T A A A A A A A A A A A A A A	ILE LYS GIN VAL A T T A A A C A G G T T G 1160GLU GLY THR ASP SER A G G G T A C C G A T T C G C 1170	ILE ILE PHE GLU ATAATTTTGAAG 1220
SE T	ASN A A	ILE A T	ASN ILE AACATA
ALA ILE ILE CAATTATA? 1030	ALA SER GLY SER THR	ASN ILE A A T A T A C 1150	VAL ALA GLU GIN TAGCCGAACAA 1210
GLY ALA ILE ILE ASN GGCGCAATTATAAAT 1030	ALA SER GLY SER THR GCTTCAGGCTCAACA 1090	GLY GLY ASN ILE GLU GGAGGTAATATAGAA 1150	VAL ALA GLU GLN ASN GTAGCCGAACAAAAC

	14	3	/	235
--	----	---	---	-----

THE INS GLY ASN ILE THR IEU GLY SER GLN IVS AIA GC C T G G C T C C A A A A B G C C C T G G C T C C C A A A A B G C C C T G G C T C C C A A A A B G C C C T C C C A A A A B G C C C T C C C A A A A B G C C C T C C C A A A A B G C C C A T C C C T T G C C T C C C A A A B G C C C A T C T C C C T C C C A C C C C C
THE GLU IIIE INS GLY ASP VAL THR C C A A C A G A A T A A A G G C G A T G T T A C C G 1270 1280 VAL LYS GLY GLY GLY VAL LYS GLY GLY T C A A A C A G G 1330 ALA ASN PHE ASP ASN HIS LYS GLY ALA T C A A A G G C G T A C C T 1330 ASN THE ALA ASP GLY ASP THR IIE ASN THR IIE NAL ASN T A A T G T G A A C T A C G G G G G G G G G G G G G G G G G G

144/235

ASIN

GAC

AAT

ASIN

ACCAACAAC 泯 盟 A C ACCTTT AATGTT 出 ASIN CTGGC GGT GLY ⊱ 開 园 TAAACCT TTA ASIN A A C A... : : Z, AAC CAAASIN WAL G A A A GATATAACA .. A C C T A GGGTG 国 WAL AGCACA ...1470 ...GLY HE 讯 ...ASIN 1460 ASP SER ပ ATC GLY . G ی Ø ASIN TTTAA TCTA C T AAA LYS E ALA ည ဗ 9

GGA ر ر ر ر ARG : A T T A T C (1610 田田田 HE CAAAAAAGC 段 LYS GIN AGCGGT GLY 贸 ... A G ...LYS

TAAACATCACCACTCACGCCA...

開

ASN

H

LYS

드

AAA

G G C

AAT TTA GGGAAT ASN GLY CAA GIN AACAAA LYS ASN ⊱ ø ITE

1	۷	5	/	2	3	5
,	_	J	•	-	$\overline{}$	~

	145/235		
GLU G A A 1680	ASN A A T 1740	THR A C A 1800	ASN A A C 1860
ILE A T T	VAL G T C	SER A G C	LEU
GLU 3 A A	LYS A A A O	SER ICAA	ASP ASN LEU SACAATCTA?
ASN ALA GLU ILE A A T G C T G A A A T T (SER ASP LYS VAL	SER SER SER 1790 - 1790	ASP A G A C 7 1850
ASN A A T (SER ICT(SER FCT7	THR ZAGACAG
SER AGT/	SER FCT	SER AGT1	PHE '
ASN A A T A	ILE ATT?	GLU 5 A G 7)	THR ACAT
ASN P A A T A 1660	 3 IHR I A C C A 1720	 GLY G G G G	 A LEU 1 T T A A 1840
ILE THR ASP ASN ASN SER ASN ALA GLU ILE GLU TTACGGACAATAATAGTAATGCTGAAATTGAA 1650 1660 1680	G CAAAAAGAAG G CAAAAAAGAAG 1700GLY ASN LEU THR ILE SER SER ASP LYS VAL ASNGTAATCTCACCATTCTGATAAAGTCAAT 1740GTAATCTCACCATTTTTTTTTTTTTTTTTTTTTTTTT	⊲' ⊑	LYS A A A GLU GLU
IR A C G (GLY GLY ASN ILE SER GLN LYS GLU GCGCAATATCTCGCAAAAGAA 1690GLY ASN LEUGTAATCTCGTAATCTC	ILE THR ILE LYS ALA TAACAATCAAAGC 1760GLY VAL ASPGCGTTGA7	ILE ATTA LYS AAA(
ILE T T 1650	E GLN 1700GLYGTY	R ILE 1760 GLY G C G	U THR A A C CTHR C C 7
: : :	SER T C G 1	THR ACA 1	LEU CTA. 1
	ILE A T C	ILE A T A	ASN A A T
	ASN A A T O	THR LYS GIN CCAAACAGA 1750	AIA G C C
	GLY AG G G G C A 1690	LYS (A A A A C 1750	ASP P G A T G 1810
		ILE THR LYS GIN ILE THE ATTACCAAACAGATAAC 1750	ALA SER ASP ALA ASN LEU THR ILE GCAAGTGATGCCAATCTAACCATTA 1810 1820THR LYSCCAAGG
	ILE ATTC	ILE A T T A	ALA G C A

ASN AA ⊱

SE E

	1407233		
GIN G G C 1920	SER T C A 1980	SER A G C 2040	
AGCAGAAATTA 1880THR ALA LYS ASP ASN SER ASP LEU ILE ILE GLYCAGCTAAAGATAACAGTGATTTAATTGGC1890 1900	SER ASN ALA CAGTAATGCTA 1940LYS GIN VAL THR PHE ASP LYS VAL LYS ASP SERAACAAGTAACCTTTGACAAGGTTAAAGATTCA 1980	TCACAATGTAA 2000THR LEU ASN SER LYS VAL GLU THR SER ASN SERCACTAAATAGCAAAGTGGAAACGTCTAATAGC	
ILE A T T A	LYS A A A G	SER	
ASP LEU 3 A T T T A A	ASP LYS VAL 3 A C A A G G T T A	HIS ASN VAL TCACAATGTAA 2000THR LEU ASN SER LYS VAL GLU THRCACTAAATAGCAAAGTGAAACGTCAOTAAATAGCAAAGTGGAAACGT	
ASP GAT	LYS A A G	GLU GAA	
SAAATTA ALA LYS ASP ASN SER CTAAAGATAACAGTO	ASP G A C	VAL G T G	
 ASN ATAAC 1900	·· ·· ·· PHE ·· C T T T C	LYS A A A A 20	
 ASP G A T 190	CAGTAATGCTA 1940 LYS GIN VAL THR AACAAGTAACC	 A SER I A G C A	: : : : : :
AAGCAGAAATTA 1880THR ALA LYS ASPCAGCTAAAGA	SN. SER ASN ALA A C A G T A A T G C T A 1940LYS GLN VAL THR A A C A A G T A A C 1950	SN HIS ASN VAL A T C A C A A T G T A A 2000THR LEU ASN SERC A C T A A A T A G	SN GLY SER ASP ACGGTAGCGATG
AGA ALA ALA	ASN SER ASN A C A G T A A T 1940LYS GIN A A C A A1950	ASN HIS ASN VAL A T C A C A A T G T A 2000THR LEU ASN C A C T A A A T	ASN GLY SER ASP AACGGTAGCGA7 2060
A G C A G 1880 THR C A G	SER C A G T 1940 LYS A A	T C A C 2000THR C A C A C A	GLY 2 G G 7 2060
A	ASIN A	ASIN A	< 1
TAAT	ASP TGAC	TGGC	THR GLY ACCGGA
1870 1870	R SER 3 C A G T (1930	rr ala 2 a g c t g 1990	R THR CAC 2050
A G G	ALA SER ; C T A G C A 1930	ILE SER TCTCA(GLY SER 3 G T A G C A 2050
ATTTCAGGTTTTAATA	LYS ALA SER SER ASP 1 A A G G C T A G C A G T G A C A 1930	LYS ILE SER ALA GLY AAAAA AAAAAAAAAAAAAAAAAAAAAAAAA	ASP GLY SER THR GLY AG ATGGTAGCACCGGAA
ATT	LYS A A	LYS A A	ASI G A

•	~	•
L	<u>マ</u> ハ	5
	· r	•
_	1	-

	147/235		
LYS A A A 2100	GLU G A A 2160	VAL G T A 2220	ALA G C A 2280
ALA G C A A	SER TCA(GLU GAA	THR ACAC
SER I C C	ALA G C A 1	SER VAL GLU GCGTGGAA	SER VAL GTGTTP
THR ILE SER CTATTTCC (2090	ILE SER TCTCTG 2150	SER FAGC 2210	SER A'A G T 2270
THR ACT	ILE ATC 2	GLY G G T	VAL 3 T A S
ASP ASN ASN ILE GLY LEU THR ILE SER ALA LYSACAACAATATCGGCTTAACTATTCCGCAAAA2070 2080 2100	TAATATCACCT 1 A A T A T C A C C T 2120SER HIS LYS THR VAL ASN ILE SER ALA SER GLUCTCACAAACAGTAAATATCTCTGCATCAGAAC130 2140	S ALA GLY THR A G C A G G C A C A A 2180THR ILE ASN ALA THR THR GLY SER VAL GLU VALC C A T T A A T G C G A C C A C G G T A G C G T G G A G T A 2220	P ILE SER GLY TATTAGCGGTA 2240THR ILE SER GLY LYS THR VAL SER VAL THR ALACGATTCCGGTAAGACAGTA'AGTGTTACAGCACGATTTCCGGTAAGACAGTA'AGTGTTACAGCA
ASP ASN ASN ILE GLYACAACAATATCGGCT2070	TAATATCACCT 2120 SER HIS LYS THR VALCTCACAAAACAGTAA	THR ACC.	LYS A A G A
ILE (A T C G 2080	 THR V A C A G 2140	 ALA 1 G C G A 2200	 A GLY 1 G G T A 2260
ASN A A T	THR ACC LYS AAA	THR ACA ASN AAT	SER GLY GCGGT ILE SER TTTCC
ASN A A C	ILE A T C A HIS C A C A	ALA GLY A G C A G G C A 2180THR ILE C C A T T A2190	P ILE SER TATTAGC (2240 THR ILE CGATTT
A C	ER ASN STAATA 2120SER CTC	S ALA A G C A 2180 THR C C	ALA LYS THR GLY ASP ILE CTAAAACAGGCGATATTA 2230 2240THRCGA
: : :	SER AGT 	LYS A A A A	ASP G A T
	ASN A A T	THR ACT	C G G
	VAL THR VAL ASN SER TAACGGTAAATAG7 2110	THR ACT 70	THR ACA 30
	THR VACGG 2110	ILE 1 A T C A 2170	LYS 7 A A A A A 2230
	ASP VAL THR VAL ASN SER ASN ILE THR GATGTAACGGTAAATAGTAATATCACCT 2120SER HIS LYS THRCTCACAAAAC	GLY GLY ILE THR THR LYS ALA GLY THR GGAGGTATCACTACTAAAGCAGGCACAA 2170 2180THR ILE ASN ALACCATTAATGC	THR ALA LYS THR GLY ASP ILE SER GLY ACTGCTAAAACAGGCGATATTAGCGGTA 2230 2240THR ILE SER GLYCGATTCCGG'
	ASP GATG	GLY G G A	THR ACT G

A A A T A T C *F* 2690

ATAAATGGGT 2680

... TAAGCACA

Ø

7	4	9	/	2	3	5

149/235		
GLY G G T 2580	THR A C C 2640	SER
SER TCA	LYS A A A	ILE
ALA G C A	ALA G C G	ILE
THR A C G 2570	THR A C T 2630	ASN
GLY G G T	VAL GTG	LEU
ASP GAT	SER AGC	GLY
LEU CTA 60	GLY GGT	ASIN
 A LYS A A G	 G SER T C T 266	$\begin{array}{c} \dots \\ \text{I} \\ \dots \\ \text{II.E} \end{array}$
ALA G C A ALA G C T	ASN A A C GLY G G T	ASP GAT THR
ASN A A C ASP G A C	THR ACT SER AGT	GLY G G G SER
ILE A T T 2540 LYS A A	ALA G C A 2600 ALA C A	3 THR GLY ASP C A C C G G G G A T T 2660LEU SER THR ILE
LEU TTA	VAL G T A	ASN A A T
THR ACC	GLU GAA 30	VAL GTG
GLY G G C 253	THR A C A 255	ASN V A A T G 2650
SER A G T	ARG C G T	SER SER ASN VAL ASN ILE TCAAGTAATGTGAATAT 2650
THR A C C	ASN A A C	SER T C A A
	fl C	

T T C 2880

G

Ø

T G

TGC

TGTAAG

TTGG

 \mathcal{O}

AAA

 \mathcal{O}

... T A G C

...2850

ATT

ACG

ILE

ALA

ASIN

PR0

VAL

GCCATT? 2900

ASN A A T

3 C C A A A T A 2890

GAG

GTT

				,,,,									
		GIN	CAA	7			LYS	AAA	2820				盟
		ARG GLY LYS GLU ILE ASP VAL LYS TYR ILE GIN	G A G G C A A G G A A A T T G A T G T G A A A T A T				ARG VAL LEU GLU LYS VAL	TTGAAGCGAAACGCGTCCTTGAGAAGTAAAA					ARG
		TYR	T A T				LYS	AAA					VAL
		LYS	A A A A				GEO	GAG	2810				ALA
		VAL	G T G	•			DET!	СТТ	(7				SER
		ASP	G A T				VAL	GTC					VAL
		H	ATT	>			ARG	C G C	00				GLY
:	Α	::OTB	GAAA	3	: A	:	LYS	AAA	2800	:	: ::	:	LEU
LEU	CTGTGCGCTTAA	LYS	AAG		AL GIU GIU VAL TAGAAGAGGTAA		ILE GLU ALA LYS	SCG			AAAGAGAAACAC		LEU ALA LYS LEU
ARG	0 9 0	GLY	S B S		GLU GAG		GLU	GAA		CIN	GAA		ALA
THR VAL ARG LEU	FGTG	ARG	G A		VAL GLU GLU VAL 3TAGAAGAGGT?	2780	. ILE	T T	2790	GLU ARG GLU THR	A G A	2840	LEU
		•	•				•	•	•			•	:
GLU ASN GLY ARG ASN	GAAAATGGTAGAAACA				PROGLY VAL ALA SER V CCAGGTGTAGCAAGCG					ASP GLU	GATTTATCTGACGAAG		
ARG	STAGA	9			AIA G C A	2770				ASP	GAC	2830	
GLY	1 G G T	3			VAL GTAG	27				SER	TCT	78	
ASN	AAAT				GLY GGTG					ASP LEU	TTA		
GLU	GA 1				PRO C C A G					ASP	GAT		

(_)
L	()
C	1	ĺ
(_	5
Ē	Ī	_

	151 / 235	
SER T C A 2940	VAL G T A 3000	
PRO C C G T	ALA G C A	
.ASN THR GIN ASN GLU PHE THR THR ARG PRO SER .ATACACAAATGAATTTACAACCAGACGTCA .2910 2920 2930	GLU GLY LYS GAAGGTAAGG 1960 ALA CYS PHE SER GLY ASN GLY ALA VAL . CGTGTTTCTCAAGTGGTAATGGCGCAGTA . 2970 . 2970	
THR A A C C 2930	GLY F G G C 2990	
THR ACA 2	ASN AAT	
PHE T T T	GLY GGT	
GLU GAA 20	SER AGT	·
ASN G A A T G 2920	GLU GLY LYS G A A G G T A A G G 2960ALA CYS PHE SER S C G T G T T T C T C A A C C 2980	: ; :
GLN C A A	GLU GLY LYS 3 A A G G T A A G G 960 ALA CYS PHE SER C G T G T T C T C C 2970	ASP GLY GLN 5 A T G G A C A G (120 PRO *** C G T A G
THR A C A	GLY GGT CYS TGT	ASP GLY G A T G G A 1020 .PRO *** . C G T A G
ASN A T <i>P</i>	GLU C A A 2960 ALA C G'.	ASP GLY G A T G G , 3020PRO *** C G T A C
• • •	SER TOT	ASP G A C
	ILE ATT.	ALA GCT
	L ILE GATA 2950	SN VAL A T G T T 3010
	VAL GTG 29	ASN A A T 30
	SER GIN VAL II.E II.E SER GLU GLY LYS A G T C A A G T A A T T T C T G A A G G T A A G G 2950 2950 2950ALA CYS PHE SERC G T G T T T C T C C	CYS THR ASN VAL ALA ASP ASP GLY GLN IGTACCAATGCTGCTGACGATGGACAGC 3010 3020 ***PRO ***CGTAG
	SER A G I	CYS I G I

Strain 15 hmw1A sequence	LYS GLU TRP LEU LEU ASP PRO ASP ASN VAL A A A G A G T G G T T A G A C C C G G A T A A T G T A A 30 10 THR ILE GLU ALA PRO SER TYR SER ARG GLYCAATTGAAGCCCCTTCCTATTCTCGCGGT 60	ASN ALA GLY ILE ASP SER GLU PHE PRO GLY A A T G C C G G C C C 90 70GLY SER GLY THR LYS GLU SER PRO LYS THR 110G T T C G G C C C C A A G C C C T A A A C A C C C C C C C C C C C C C C C C C C C C	ASN GLY GLU GLA PRO THR VAL LEU THR ASN A A C G G C G A C A G T A T T A A C C A A T G 130 1.GLU THR ILE SER ASN TYR LEU LYS SER GLY 1.CLU THR ILE SER ASN TYR LEU LYS SER GLY
Strain	LYS	ASN	ASN
	A A A O	A A T G	AACO

SUBSTITUTE SHEET (RULE 26)

LE T A 420	LA C G 480	LE T T 540	T T
ILE ATA 42(ALA S G C (ILE AT 54	PHE
SER T C C A	THRACCO	SER T C T	ARG A G A
LYS GLY AAAGGA1	ILE A T C A	VAL 3 T A	HIS
LYS A A A (410	GIN C A A A 470	ASN VAL A A T G T A 1 530	SER
ALA 3 C T P	ALA 5 C T C	ASN AATA	THR CTT
THR A C A G	···· S語 SBR	 LEU T T A	 T ASN A A C A
ILE TT2	ALA CCC 450 SER CAP	CAGGGAACTATAACCAATACTGGCGATCAAA CAGGGAACTATAACCAATACTGGCGATCAAA 490LYS GIN LEU ARG LEU ASN VAL SER ILEAACAACTCAGACTTAATAATGTATCTATT 520 530	VAL TT7 570 PRO CT7
ASN] A A T A 400	LYS A A A G ALA 3 C A T 460	ASP GIN S A T C A A 5.10 LEU ARG T C A G A 520	PHE T T T T T T T T T T T T T T T T T T T
LEU PTGA	GLU G A A A ASN ,	GLY 3 G C G GIN CIN	LEU ASN FTAAATT 560 SER ILE CAATTC
C C T	THR TACAG440 GCA	1 THR GLY TACTGGC (500LYS GLN AACAAC	LEU T T A A 560 .SER . C A A
	GLY 5 G T 7 7	ASN A A T A E	GLY 1. G T T 5. C T T
	ASN A A C G	THR ACCA	ILE GLY TAGGTT
	GLY ASN GGAAAC O	ILE THR ATAACC 0	GLY ; G T ?
	GLU CG A G G 430	THR I CTA 490	TAR (C G G 550
	PHE L T T G	GLY THR 3GAACT 49	ASN GLY THR GLY ILE A A T G G G A C G G G T A T A 550
	AIA ; C C T	GLN 'AGG	ASN A T G

FIRE GLY ARG VAL 620 630 HIS VAL ASN GLN THR THR PRO LYS ASN LEU A T G T T A A T C A A A C C A C A C A C A C T G ASP GLU SER TYR C G A T G A A T C T T A T T 680 680 680 690 GRA T G A A T C T T A T T 680 680 680 690 GRA T G A A T C T T A T T 680 680 690 680 690 680 690 680 690
--

|--|

•	157/235		
VAL SER ASP GLY SER ASN LEU THR PHE ASN GTTTCTGATGGCTCAAACCTTTTTAATT 1030SER SER ILE ARG GLY GIN GLU ALA PHE ASNSER SER ILE ARG GLY GIN GLU ALA PHE ASNCTTCCATTCGCGGCCAGGAAGCCTTTAATCTTCCATTCGCGGCCAGGAAGCCTTTAAT	ILE SER LYS ASP LEU THR ILE ASN ALA THR A T C A G A T T T A A C C A T A A A T G C A A C C G 1100 1100GLY SER PHE PHE GLU LEU GLY GLN TYR SERG T T C A T T T T T G A A C T T G G C A A T A C T C GG T T C A T T T T T T G A A C T T G G C A A T A C T C G	ASP THR PHE ASN GLY ASN GLY PHE ASN HIS GATACCTTTAATGGTAATGGCTTTTAACCACG 1160 1170 1.ASP ALA ILE LYS SER THR HIS ASN ILE SERASP ALA ILE LYS ATATATATCA 1180 1190 1200	ILE LEU GLY GLY ASN VAL THR LEU GLY GLY ATCTTAGGTGGCAATGTTACCCTTGGCGGGC 1210 1220 1230

7. 50 7. 50		GIN ASI	ASP SER SER SER 3 A T T C A A G C A G T A 1240	GIN ASP SER SER THR ILE THR GLY ASN A A G A T T C A A G C A G T A C C A T T A C A G G T A T T A C A G T A A T 1240 1250 1260
ILE ASN ILE SER GIA ATCAATATCTCTCA 1270	ASN ILE SER GIN ATATCTCTCAG 1270	ALA G C	N ALA ALA ASN VAL THR GGCAGCAAATGTTACCT 1280 1290LEU ARG ALA TYR ASNTGCGAGCTTATAATC	AGCAAATGTTACCT 1280 1290LEU ARG ALA TYR ASN GLY AS
LYS GLN LA AAACAAC	LYS GLN LEU THR LEU AACAACTAACCCTT 1330	GLY ASN VAL TGGCAATGTAT 1340GLU GLY AAGGGA	LYS GLN LEU THR LEU GLY ASN VAL SER ILE A A A C A A C T A A C C C T T G G C A A T G T A T C T A T T G 1330 1340 1350GLU GLY ASN LEU SER A A G G G A A T T T A A G T T	CAATGTATCTATTG 1340 1GLU GLY ASN LEU SER LEU ILE GLY ALA SERAAGGGAATTTAAGTTTAATCGGTGCAAGTAAGGGAATTTAAGTTTAATCGGTGCAAGT
ALA ASN ILE ASN GLY GCAAATATTAACGG 1390	LE ASN GLY ITAACGG(1390	ASN GLY ASN LEU SE ACGGCAACCTTTC 1400 GLU AST	Y ASN LEU SER VAL LYS CAACCTTTCCGTTAAAG 1400 1410GLU ASN ALA LYS PHEAAAATGCTAAATTT?	CCTTTCCGTTAAAG 1400 1400GLU ASN ALA LYS PHE LYS GLY GLU THR GINAAAATGCTAAATTTAAAGGGGAAACCCAA 1420 1420 1420

... TAAATATTCAGGTTTCAATAAAGCAGAG ... 1840 1850 1860

LYS A A A 1680	LYS A A A 1740	ILE A T T 1800	GLU
GLN C A A	ILE A T C A	THR A C C	ALA
SER T C G C	THR ACAA	LEU CTA	LYS
1 ILE TATC 1670	ILE 3 A T A A 1730	ASN 2 A A T 1790	ASN
TLE GIN ILE GLY GLY ASN ILE SER GLN LYSTCCAAATTGGCGGCAATATCTCGCAAAA 1660 1670 1670	GIN	GAATTCCGATTCAG 1760 1760ASP ALA THR ASN ASN LEU THR ILEASP ALA THR ASN ASN ASN LEU THR ILEACGCGACAATGCCAATCTAACCATTACGCGACAAACAATGCCAATT	PHE
G G C	LYS A A A A 1710 THR ASN C C A A T	G ASN A A T	 C GLY
G G C	LYS A A A A 1710 THR A C C A	SER T C A C 1770 ASN A A C A	THR GIN ASP ACGCAAGACC 1830 ASN ILE SER GLY
ILE GI ATTG(1660	ASP 1 GATA IIE TAATTA	ASP G A T T T THR PACAA	GLN C A A I
GIN C A A	SER TCC ASN ASN	SER T C C ALA 3 C G 7	THR A C G A
. T C	HR II.E SER SER ASP LYS CGATTATCTTCCGATAAA 1700 1710 II.E ASN II.E THR ASNTCAATATTACCAA	ASP GLY GLU ASN SER ASP SER SATGGGAGAATTCCGATTCA 1760 1770ASP ALA THR ASNACGCGACAAAC	LYS LEU THR GIN ASP AATTAACGCAAGAC 1820 1830 ILE SER
: : :	II E	GLU G A G 1	LYS A A A '
	LEU THR TCACGA	GLY	LEU TTG,
	LEU CTC O	ASP G A T O	GLU G A A O
	ASN I AATC 1690	VAL <i>P</i> G T T G 1750	LYS GJ A A A G <i>i</i> 1810
	LU GLY ASN LEU TE A A G G C A A T C T C A (LA GLY VAL ASP GLY GLU ASN SER ASP SER C A G G T G T T G G G G A G A T T C C G A T T C A G 1750ASP ALA THR ASN ASNASP ALA THR ASN ASNAC G C G A C A C A C	AS THR LYS GLU LEU LYS LEU THR GLN ASP AAACCAAAGAATTGAAATTAACGCAAGACC 1810 1820 1830
	alu A A	LA C A	YS A A

III

<u> </u>	<	_
Ī	C)
C	1	1
(Γ	5
ī	ī	-

	•		
1LE A T T 2100	ALA G C A 2160	GLY G G T 2220	VAL G T A 2280
ASN AACA	ASP G A T	LYS A A A	ASN A A T
LYS ASN A A A A A C P	LYS ALA A A A G C G C	ILE A T T A	LEU
LYS A A A A 2090	LYS 7 A A A 2150	ASP F G A T A 2210	THR 2270
VAL GTAA 20	THR ACCA 21	GLY GGTC	ASP GACA
ALA LYS ASN VAL GLU VAL LYS ASN ASN ILE C G A A A A T G T A G A A G T A A A A A C A A C T T 2080 2100	ASN ILE THR ALA A A A T A T C A C C G C G T 2120 2120SER GLU LYS LEU THR THR LYS ALA ASP ALA C A G A A A A C T T A C C A C C A A A G C G A T G C A 2140 2150	GLY ASN VAL GLU T.G.G.T.A.A.C.G.T.A.G.A.B 2180VAL THR ALA LYS THR GLY ASP ILE LYS GLYT.G.A.C.A.A.A.C.A.G.G.T.G.A.T.T.A.A.A.G.G.T.	CGGTAATGTAATA CGGTAATGTAAATA 2240 ILE THR ALA ASN GLY ASP THR LEU ASN VAL TTACAGCAAACGGCGACACGCTTAATGTA TTACAGCAAACGGCGACACGCTTAATGTA
N VAL ATGTA 2080	VAL ASN ILE THR ALA G T A A A T A T C A C C G C G T 2120 2130 SER GLU LYS LEU THR C A G A A A A A C T T A C	THR GLY ASN VAL GLU A C T G G T A A C G T A G A A G 2180 2180VAL THR ALA LYS THRT G A C A G C C A A A A C C 2200	SER GLY ASN VAL ASN T C C G G T A A T G T A A A T A 2240 2250ILE THR ALA ASN GLY T T A C A G C A A A C G G G
ASN A A T	THR A C C C LYS A A A C 2140	VAL GTAC ALA GCCA 2200	ASN VAL A A T G T A A THR ALA C A G C A A 2260
LYS A A A	ILE ATCA GLU GAAA	ASN A A C THR A C A	ASN A A T THR A C A
ALA	VAL ASN TAAATP 2120SER CAG	THR GLY ASN A C T G G T A A C G 2180VAL THRT G A C A G	SER GLY ASN VAL ASN T C C G G T A A T G T A A A T 2240 ILE THR ALA ASN T T A C A G C A A A C 2260
	VAL G T	A C '	SER T C
	LYS THR AAAACA)	THR A C C	SER THR
	N LYS	ASN ALA THR AATGCAACO 2170	S SER A T C C 2230
	ASN TAACA 2110	ASN TAAT	VAL LYS 3 T C A A A T 2230
	THR SER ASN LYS THR ACTTCTAACAAAACA 2110	THR ILE ASN ALA THR ACCATTAATGCAACC 2170	GLU VAL LYS SER THR GAAGTCAAATCCACT 2230
	THR A C '	THR A C (GLU G A A G

LEU THR THR GIN	ALA LYS ASP GLY SER	THR LEU THR THR	
TTAACCACCCAA	CCAAGGATGGCAGT	ACTTTAACTACT	
2330 2340	2390 2400	450 2460	
SER ASN VAL SER GLY ASN ALA VAL THR ILE AGTA ATGTTTCAGGCA ATGCTGTTACCATCA 2310 2290THR ALA ASP LYS GLY LYS LEU THR THR GINCTGCAGATAAGGGCAAATTAACCACCCAA 2340	ALA SER SER ILE THR SER ASN ASN GLY GCAAGCTCTAGCATTACCTCAAACAATGGCC 2350 2360 2370GIN THR THR LEU THR ALA LYS ASP GLY SER AGACAACTCTTACAGCCAAGGATGGCAGT	ALA GLY SER ILE ASN ALA ALA ASN VAL CAGGAAGCATCAATGCCGCCAATGTGA 2420 2430THR LEU ASN THR THR GLYCATTAAATACCACAGGCA 2440	VAL GIU GLY SER ASN ILE ASN ALA ALA SER GTAGAAGGTTCAAACATTAACGCAGCCAGTG 2470 2480
SER A	ALA	ILE	VAL
	G C A	A T C G	G T A

ALA G C T 2520	GLY G G C 2580	THR A C A 2640	THR ACT 2700
ASP GAT(SER A G T	SER A G T	GLY G G T
ALA LYS ; C A A A A G 10	ALA 3 C A A	LEU	LYS A A A (
ALA F G C A 7 2510	ASN F A A C (2570	ASP 3 G A T 1 2630	LEU 3 T T A 7 2690
ASN A T G	THR 1 C T A 255	GLY 3 G G G	VAL LEU ; T G T T A 7
ILE \ T T A	 A AIA G C A A	 A THR ACCG	 VAL
GLY THR LEU VAL ILE ASN ALA LYS ASP ALA GTACCTTGGTTATTAATGCAAAGATGCT GTACCTTGGTTATTAATGCAAAGATGCT	HIS A C 2550 2550 ASN A T (GGTAACCTCAAGTA GGTAACCTCAAGTA 2600 2.000ASN VAL ASN ILE THR GLY ASP LEU SER THRASN VAL ASN ILE THR GLY ASP LEU SER THRATGTGAATATCACCGGGGATTTAAGTACA 2620 2640	CATTTCGAAAATG 2660 2670 CLATTTCGAAAATG 2660 2670 CLY ARG ASN THR VAL VAL LEU LYS GLY THRGLY ARG ASN THR VAL VAL LEU LYS GLY THRGLY ARG ASN THR VAL VAL LEU LYS GLY THR
LEU TGG	GLY ASP] S G T G A C C VAL VAL A S T A G T A A 2560	SER CCAA ASN ASN 2620	LYS A A A A ASN A C A 2680
THR CCT	GLY 3 G T C VAL 3 T A G	THR SER ACCTCA VAL ASN TGAAT	SER LYS L C G A A A ARG ASN G A A A C A G A A A C
GLY G T A	SER A T C A C 2540 THR C A G	G G T A A 2600ASN	CATTT CATTT 2660 GLY
	ALA , C	VAL THR ALA VAL THR SER SER STGACTGCGTAACCTCAAG 2600 ASN VAL ASN ILEATGTGAATATC 2620	ILE 1 T C A 26 26
	ALA 5 C G (THR A C T O	ASN A T P
	GLY 5 G C (VAL 3 T G A	LEU TA7
	ASN GA CG 2530	SER 1	GLY I 3 G A T 2650
	JYS LEU ASN GLY ALA AGTTGAACGGCGCGG	SER GLY SER VAL THR ALA VAL THR SER SER C T G G T A G T G A G T A 2590 2600 2.ASN VAL ASN ILE THRASN VAL ASN ILE THR 2620	AL ASN GLY LEU ASN ILE ILE SER LYS ASN TAAATGATTAAATATCATTTCGAAAATG 2650 2670 CLY ARG ASN THR VALGTAGAAACCGT
	LYS A G .	SER CT(AL T A I

-1G.26N

	ALA G C G 2760		LYS A A A 2820		GLN	C A A	7007
	GLU GAA		ALA G C T		强	A C A	
	ILE A T T		LEU T T A		ASN	AAC	
	NAL A G T A 2750		A A C A 2810		VAL	G G T T	
	VAL ALA SER VAL GLU GLU VAL ILE GLU ALA TAGCAAGTAGAAGAAGTAATTGAAGCG 2740 2750 2760		SER ASP GLU GLU ARG GLU THR LEU ALA LYS C T G A T G A A G A A G A A A C A T T A G C T A A A A C A A C A T T A G C T A A A C A C A T T A G C T A A A C A C A T T A G C T A A A C A C A T T A G C T A A A C A C A T T A G C T A A A C A C A T T A G C T A A A C A C A T T A G C T A A A C A C A T T A G C T A A A C A C A T T A G C T A A A C A C A C A C A C A C A C A C		PRO ASN ASN THR ILE THR VAL ASN THR GIN	CAAATAACCATTACGGTTAACACACAA	1
: : : : : :	G A A	: : :	ARG A G A	: :: :	ILE	ATT	
GLY G G T 2730	VAL GTA	LEU T T A 2790	GLU GAA 10	GLU GAA(2850	THR	ACC	2
PRO C C A	SER V7 AGTG7 2740	ASP GAT	GLU CG A A G 2800	ILE A T T	ASN	A A T A A	707
GIN C A G	ALA G C A	LYS A A A	ASP G A T	PHE T T T	ASIN	AAT	
LYS TYR ILE GIN PRO GLY AAATATATCCAGCCAGGT 2720	. T A	VAL A G T G 2780	SER C T	ARG A C G T 2840	. PRO	. C A	
TYR TAT	: : :	LYS A A A	: : :	VAL G T A	:	:	
LYS A A A		GLU GAG		ALA G C T			
ILE GLU VAL \TTGAGGTGA 2710		AL LEU FCCTT 2770		SER A G T 30			
GLU GAG 27.		VAL G T C 27'		VAL S G T A A 2830			
GLU ILE GLU VAL LYS TYR ILE GLN PRO GLY GAAATTGAGGTGAAATATATCCAGCCAGGTG 2710 2720		LYS ARG VAL LEU GLU LYS VAL LYS ASP LEU A A A C G C G T C C T G A A A G T G A A A G A T T T A T 2770 2780		LEU GLY VAL SER ALA VAL ARG PHE ILE GLU CTTGGTGAGTGCTGTACGTTTTATTGAAC 2830 2840 2850			
GLU GAAA		LYS A A A		LEU CTT			

SUBSTITUTE SHEET (RULE 26)

FIG.26C

TGT A G 3 C C GGACAGCAG GGTAAG LYS GAB GTTGCTGACGAT ASP AGTCAAG... 2910 ATTTCT ASP SER SEC TCA GGTAATGGCGCAGCAGTA SE VAL TTTACAACCAGACCA ... T G ... A T ...VAL ... ASIN ARG ALA 111 强 ASIN AATGAG GLU SER T C A ASN SER

166/235

FIG.27/

NTHi strain 15 hmw2A sequence

167/235		
GLY G G 60	THR A C 120	LEU T T 180
ALA G C A	ASP G A T	ILE ATAT
ASP G A C	ALA G C T	LYS A A G
GLY . G G A 50	SER AGT	GLU GAG
ALA G C A ·	GLU S GAGA 110	LEU GCTTG
SER HIS MET PRO CCCATATGCCG 20 GLU ASN VAL TYR ILE ASN ALA GLY ASP ALA GLY GAGAATGTATATATTAATGCAGGAGACGCAGG GAGAATGTATATATTAGCAGGAGACGCAGG 30	LEU GLU ASN GLU TAGAAAACGAA 80 GLU TYR THR GLY GLU SER ALA ASP THR GAATACACAGGAACAGGAGAGTGCTGATAC GAATACACAGGAACAGGAGAGTGCTGATAC 90	ASN THR LYS THR ACACAAAGACA 140 THR LEU THR ASN SER THR LEU GLU LYS ILE LET ACACTAACTCAACGCTTGAGAAGATATT ACACTAACAAACTCAACGCTTGAGAAGATATT 150
ILE A T T	THRACA	SER T C A .
 TYR T A T	 GGA	 ASN A A C '
PRO . C C G. VAL G T A .	GLU . G A A. THR .	THR. ACA. THR
MET A T G ASN A A T	ASN A A C TYR T A C	LYS A A G LEU C T A
SER HIS MET PRO 20 GLU ASN VAL T GAGAATGTAT 30	TAGAAAACGAA 80 GLU TYR THR GGAATACACAG	ASN THR LYS THR 1 A C A C A A A G A C A 140 THR LEU THR A A C A C T A A C A A A
SER HIS T C C C A 20 GLU G A (LEU GLU T A G A 80 GLU G A	ASN THE A A C A C 140 THR A C
GLY GGA	ASN A A T	ASN A A T
SER T C G	THR ACT	ASN A A C
ALA G C T 10	ASP G A C 70	ARG C G A 7 130
ASN SER ALA SER GLY SGANTCGGCAT	ARG SER ASP THR ASN I GCGTAGTGACACTAATT 70	PRO LYS ARG ASN ASN A TCCAAAACGAAACAATA 130
ASN A A T	ARG C G T	PRO C C A
Ö	Ŋ	£-I

FIG.27B

	100/250		
ILE A T 240	VAL G T 300	GLY G G 360	
ASP GATA	C C C	SER AGCG	
SER AGTG	SER A G T	SER T C T A	
ASN A A T	ASN A A C	TYR I A C	
VAL A 3 T T A 230	LYS 7 1 A A A A 290	ILE SATT 350	
ARG VAL ASN 1GAGTTAATA 230	ASN	THR ILE TYR ACCATTTACT 350	
ILE A T C A	LYS ASN LYS ASN SER GLY A A A A T A A A A A C A G T G G C C 290	LEU	
THR A C T ASN GLU A A T G A A A 220	SER HIS LEU THR TCCCACCTAACC 260 LEU TRP SER LYS ASN LYS ASN SER GLY VAI CTCTGGAGCAAAAAAAAAAGTGGCGT 270	ILE THR SER THR A T C A C T T C T A C T 320 ALA ASN GLY ASN LEU THR ILE TYR SER GLY G C T A A C G A A A C T T A A C C A T T T A C T C T A G C G G 350 330	LYS ASN ILE THR A A A A A T A T T A C G 380
TATC TATC ASN CAAC	C C T TRI	TTCT TTCT ASN TAAC	ILE TATI
T T A A T A 200 G C C A 210	SER HIS CCCAC 260 LEU CTC	ILE THR T C A C T 320 ALA G C T	LYS ASN A A A A A T 380
ى د ك	O3	A	
PHETITI	ASN A A C	ASN A A T	HIS CAT
SER T C T	G G C A	G G C A	ILE A T T C
GLY G G C 7 190	GLY G G A C 250	ASN A A T G 310	ASP G A T A 370
ARG A G A C	ILE A T C G	ILE A T T A	VAL GTTG
ALA ARG GLY SER PHE AGCAAGAGGCTCTTTT 190	ASN ILE GLY GLY ASN CAATATCGGAGGCAAC 250	LEU ILE ASN GLY ASN TCTGATTAATGGCAAT 310	TRP VAL ASP ILE HIS A T G G G T T G A T A T T C A T 370

()
1	1	j
	_	5
Ī	L	-

20009			-
<i>!</i> ·	169/235		
LYS A A 420	ILE A T 480	THR A C 540	PHE T T 600
THR ACT	THR ACTP	G G G	TYR I A T
THR A C A	GLY 3 G A A	THR ACT(ASN A A C T
ILE \ T T .	GIN C A A G	GLY ; G C ?	THR CAP
ASN ILE THR THR A A C A T T A C A A C T A	GLY G 3 G T C 470	ASN G A A T G 530	VAL 590
	GIU LYS GLY GAAAAGGG 10 ASN ASN LEU THR ILE THR GLY GIN GLY THR ILE AATAACCTAACCATTACAGGTCAAGGAACTAT 450 450	GLY PHE ARG GGCTTTAGA OO PHE GLU ASN VAL SER LEU ASN GLY THR GLY THI TTTGAAAATGTCTCTTAAATGGCACTGGGAC 510 520 530	SER ARG PRO A G T A G A C C A 60 GIN LYS ASN ASN SER LEU VAL THR ASN TYR PHI C A A A A A A C A A T A G T C T C G T C A C A A C T A T T T 580 570 600
ARG LEU GCTTA1	ILE '	SER]	ER I GTC
JLY 7 G A C 400	 THR ACCA	 VAL S G T C T 520	
GLU SER GLY SAATCAGGAC	GLY G G G ILEU T	ARG A G A ASN VI	SER ARG PRO A G T A G A C C A) CIN LYS ASN ASN C A A A A A C A A T A 70 580
J S	LYS GLY A A A G G G ASN LEU ASN LEU	T A T A A A A A A A A A A A A A A A A A	ACO ACO ASA AAAAAAAAAAAAAAAAAAAAAAAAAAAA
G A	LYS AAAG ASN 1	PHE TTT GLU GAAA	ARG A G A LYS A A A
LEU CTTG 390	GLU T G A A A 440 . ASN . A A T A . 450	GLY PHE GGCTTT? 00 PHE GLU TTTGAAA	SER C A G T 560 GIN C A A Z 570
: : :	PHE T T 44	LYS (A A B B B B B B B B B B B B B B B B B	LEU C T C 56
	ALA G C C T	ASN A A T 1	ASN A A T
	VAL 3 T C C	ASN A A T	PHE T T T T
	ASP 3 A T (430	GLY 3 G C 1 490	LEU 7 T T 1 550
	GLU GLY ASP VAL ALA PHE GLU LYS GLY A G A A G G A G A T G T C G C C T T T G A A A A G G G G 440 ASN ASN LEU T A A T A A C C T A A A	THR ALA GLY ASN ASN LYS GLY PHE ARG TACAGCAGGCAATAATAAGGCTTTAGA 490 PHE GLU ASN VI TTTGAAAATG	LEU TGC
	GLU 3 A A C	THR . C A G	GLY ; G C T
	A G	E K	E E

170	0/235	_	
THR ' A C 660	ALA , G C 720	PRO C C 780	
ALA G C T	ASN A A T (ALA 3 C C C	
PRO ASN ALA CTAATGCTA	LEU . T A A	ASP GLY SER 3 A T G G C T C A G 770	
RO C T A	IIS A C I	LY G C T	
PRO F CAC 650	ILE THR HIS TAACCCACT	. т. С. 770	
T C C	A A C	AS G G A	
ILE ATTC	ILE	ALA GLU 3 C A G A G C	
MET A T G A	ASN A A T A	ALA G C A	
 SER CCA 7 640	 TRP T G G 7	 SER T C G (: : :
TE SER GLY SER TTTCAGGAAGC 620 VAL ASN ILE SER MET ILE PRO PRO ASN ALA THE GTAAATGTTCAATGATTCCACCTAATGCTAC 660 630	RG TYR LYS GLY GATACAAAGGG 680 ARG THR TYR TRP ASN ILE THR HIS LEU ASN ALA CGAACCTATTGGAATATAACCCACTTAAATGC 720	HE ASN LEU THR TTAACCTTACT 740 ILE ASP SER SER ALA GLU ASP GLY SER ALA PRO ATTGACTCCTCGGCAGAGGATGGCTCAGCCCC 780 ATTGACTCTTGGCAGAGGATGGCTCAGCCCC 780	CTTAAACGGC
GLY 3 G A A ASN A A T A	LYS A A A G THR A C C T	EU TATAS	LEU ASN GLY TTAAACGG(
SER G CAG VAL A TAA	TYR LYS FACAAAC ARG THR CGAACCT	ASN LEU A C C T T A ILE ASP T T G A C T	U A A
TTTCAGGA 620 VAL ASN GTAAAT 630	3 TYF A T A 680 ARC C G	HE ASN LEU TTAACCTTA 740 ILE ASP ATTGACT 750	C T T 800
A T A	ARG A G	五 T T T T T T T T T T T T T T T T T T T	
A A T A	SER A G C A	ASN A A C T	TYR
LEU LTA?	TYR I A C A	SER A G C A	SER AGTT
ASN GLY THR LEU ASN ILE TAATGGGACTTTAAATAT 610	SER ASN TRP TYR SER ARCAGCAG 670	SER GLU ASP SER ASN PHECTCGAAGATAGCAACTT730	LEU LEU SER SER TYR THI TCTTTTATCCAGTTATAC 790
GLY 'G G A	ASN '	GIU 3 A A G	LEU S
ASN G	SER A	SER G	LEU LI
ASI TAA	SE A A G	SE CTC	LE T C T
			_

FIG.27E

			~
	171/235		
VAL G T 840	TYR T A 900	VAL G T 960	SER 1020
ASN A A T	GIN C A A T	ASN T G	ASN . A T 1
PHE T T A	ASN A A T C	LY G G A	ILE 1 T A A
THR PACCT	A A	5 5 5 7	T A I
C A C 830	N ILE TATA 890	GLY GLY GLY SGAGGGGGG 950	VAL ILE FTAATTA 1010
THR 'ARCCA 830	THR ACTA 890	GLY GGA	VAL GTA 10
ASP GATA	CLY G G G ?	SER T C A G	G C C
THR ACA(ILE A T A G	VAL GTT	PRO C C T G
THR A C C A 820	 PRO 2 C A 2	 SER T C A (THR A C C
PHE LTC1	ASN A A C ALA PS C A C	ASN A A T ILE SE A T T T C	SER FCT GIN TH
ILE SER PHE THR THR ASP THR THR PHE ASN VAJ ATATCATTCACCACAGATACCACCTTTAATGT 810 820 830	CAACTTTAAC 860 ILE LYS ALA PRO ILE GLY THR ILE ASN GIN TYR ATCAAAGCACCAATAGGGACTATAAATCAATA . ATCAAAGCACCAATAGGGACTATAAATCAATA . 870	ATTATTCAAT 920 GIY ASN ILE SER VAL SER GLY GLY ASN VAL . GGGAACATTTCAGGAGGGGAATGT . GGGAACATTTCAGTTTCAGGAGGGGAATGT . 930 940	TTCATCCTCT 980 ASN GIN GIN THR PRO GLY VAL ILE ASN SET . ASOCA A CCCCTGGCGTAATTATAATTC A CCAGCAAACCCTGGCGTAATTATAAATTC.
ILE A T A T 810	ASN PHE A A C T T T 60 ILE LYS A T C A A A 870	A LEU PHE A T T A T T C 920 GLY ASN G G G A A C 930	SER T C A T 30 ASN (A A C C
		~	A
	LYS A A A G	TYR T A C G	ASN A A C G
	ALA 3 C A A	ASN A P T T	LEU
	ASN A A T G 850	LEU CTGA 910	ARG A G G C 970
	ASN LYS ASN ALA LYS VAITAATAAAGT 850	ASN ASN LEU ASN TYR ALA CAATAACCTGAATTACGC 910	THR PHE ARG LEU ASN ALA CACCTTCAGGCTTAACGC 970
	ASN A A T A	ASN A A T	THR A C C T
	, H	Ü	Ü

FIG.27F

17.	2/235		
LYS HIS LEU ASN ALA SER LYS GLY SER TAAACACCTTAATGCTTCAAAGGGTCG 1040 SER LEU ARG PHE GLU THR THR GLY SER THR LYS AGCTTAAGATTTGAAACTACAGGTTCAACAAA 1080 1080	IEU ILE ASN ASP LEU TAATAAATAATGATTTA 1100 THR LEU ASN ALA THR GLY GLY ASN ILE SER LEU ACTTTAAACGCCACTGGAGGCAATATATGGT11101110	IEU GIN VAL GLU GLY ILE ASP GLY MET CTTGCAGGTTGAAGGCATTGACGGGATG 1150 ILE GLY GLU GLY VAL VAL AIA LYS LYS ASN ILE ATTGGTGAAGGCGTTGTAGCTAAAAAAAAT 1200	THR PHE THR GLY ASN ILE THR PHE A A C C T T T A C T G G A A T A T C A C C T T T 1210 1220
LYS HIS LEU A A A C A C C T T 1030	VAL GLY PHE GTCGGTTTT 1090	LEU GIN VAL TTGCAGGTT 1150	THR PHE THR ACCTTTACT 1210
E	A	O	A

(5
Ĺ	
\ _	7.
<u>_</u>	<u>5</u>
L	L

20009			Terrension
	173/235		
GLY \ G G 1260	ASP HIS 3 A T C A 1320	GLY ; G G 1380	ILE AT 1440
LYS GLY A A A G G 1260	ASP 3 A T	GLY 3 G C	ALA 3 C T 7
ILE A T C A	ASN A A C G	ALA GLY	LYS A A A G
GLU AAA	PHE T T 7	THR TCCC	LEU TTA
LYS ALA ILE THR GLU AAAGCCATAACAGAAA 1240 1250	ASP PHE 3 A T T T T 7 1310	GLY ASN LEU 7 3GCAACCTTA 1370	ASN LE ATCT 1430
E A F	SER 1	7 C C	A A
II (A)	S	A P P	AI S G C
ALA G C (GLY G G 7	G G C	G G C
LYS A A A 1240	 ILE A T C 1300	 A G A G 1360	 ASN A A T A
LYS A A G A	ASN A A C LEU I	ASP G A T ASN AF ASN AF A A T A G	LEU . CTT. CTT. GLU .
GLY SER LYS LYS ALA ILE THR GLU ILE LYS GL) GGCTCCAAGAAGCCATAACAGAAATCAAAG 1230 1240 1240	LU ASN THR ASN A A A A C A C C A A C 1280 ALA THR LEU ILE GLY SER ASP PHE ASN ASP HIS G C C A C T C T T A T C G G T T C G G A T T T T A A C G A T C A C A C A C A C A C A C A C A C A	LE LYS GLY ASP TAAAAGGAGAT 1340 VAL VAL ASN ARG GLY ASN LEU THR ALA GLY GL) GTCGTCAATAGAGGCAACCTTACCGCTGGCGG 1350 1360 1380	GLY ASN LEJ GGAAATCTT 100 THR VAL GLJ ASN GLY ALA ASN LEJ LYS ALA ILJ ACCGTTGAAAATGGCGCCAATCTTAAAGCTAT ACCGTTGAAAATGGCGCCAATCTTAAAGCTAT 1410
GLY G G C T 1230	ASN A A C A D ALA S C C A 290	LYS A A A (0 VAL 5 T C G	GLY C G G A A 100 THR A C C (
::12	GLU ASN 3 A A A A 1280 ALA G C (ILE LYS A T A A A A 1340 VAL G T (GLY (3 G C G 1400 A A
	ASN VAL THR ILE ASN GLU ASN THR ASN CAATGTTACTATCAATGAAAACACCAAC 1280 ALA THR LEU IIGCCACTCTTA?	LYS LYS PRO LEU ASN ILE LYS GLY ASP TAAAAACCTTTAAATATAAAAGGAGAT 1330 VAL VAL ASN AI GTCGTCAATA(ASN VAL ILE ASN ILE GLY GLY ASN LEU CAATGTTATCAATATAGGCGGAAATCTT 1390 THR VAL GLU AACCGTTGAAA
	ILE \ T C A	LEU TA	ASN A A T A
	THR A C T A 1270	PRO C T 1	ILE A T C A 1390
	VAL ; T T ?	LYS 1 A A C	VAL
	SIN A T G	YS A A A	SN A T G
	C A Z	T A E	C A J

174/235

FIG.27H

	1747 233		
THR ASN PHE THR PHE ASN VAL GLY GLY CACAAATTTCACTTTTAATGTGGGGGC 1450 LEU PHE ASN LYS GLY ASN SER ASN ILE SER TTGTTTAACAACAAGGCAATTCAAATATCTC 1490 1500	ILE ALA ARG GLY GLY ALA LYS PHE LYS CATTGCTAGAGGGGCTAAATTTAAA 1520 ASP ILE ASN ASN THR SER SER LEU ASN ILE THR GATATCAATAACACCAGTAGCTTAAATATAC 1560	THR ASN SER ASP THR THR TYR ARG THR CACCAACTCCGACACCACTTACCGTACC 1570 ILE ILE GLU GLY ASN ILE THR ASN LYS ALA GLY ATTATAGAAGGTAATATAACCAACAAAGCAGG 1620 1620	ASP LEU ASN ILE ILE ASP ASN LYS GLY TGATTTGAATATAATAAGGT 1630 1640
Д. Т. Д. Т.	A 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	S T C	A; A i A i
ASN ASN A T	ALA GCT	ASN A A C	LEU TTG
THR CACAA	ILE CATTG	CACC	ASP TGAT

7	
9	•
Щ	•

GIN LYS GLU GLY ASN LEU THR ILE SER GLAAAAGAAGGTAACCTCACATTCCAAATTGGCGGCAACATCTC 1680 GCAAAAAGAAGGTAACCTCACGATTTCC 1690 1700 SER ASP LYS ILE ASN ILE THR LYS GIN ILE THR C	Ŋ	THR ILE LYS THR LYS GLU LEU LYS LEU AACCATTAAAACCAAAGAATTA 1810 THR GIN ASP LEU ASN ILE SER GLY PHE ASN LYS ACACAAGACCTAAATATTTCAGGCTTCAACAA 1860 1860
--	---	--

_)
[[`	į
_	· `	7
7	_	2
L	L	_

SER

VAL

AT ASP

17	76/235		
~	۵.	ASN ASP ASN ASN THR AATGATAACAACAC 2030 2040	
AAAGATAGTAGT 1880 ASN LEU THR ILE GLY ASN SER ASP SER GLY AATTTAACTATTGGTAATAGTGATGATACCGG AATTTAACTATTGGTAATAGTGATGATAG	YS ASP A A G A T T 960	THR LEU ASN SER LYS CACTAAATAGCAAA 2000 VAL LYS THR LEU SER ASP ASN ASP ASN THEGTGAAACACTTAGTGATAACACAC 2040	ASN ASN THR GLY AACAATACCGGT 2060
AGCAAAGATTGTAGCTAAAGATAGTAGT 1870ASN LEU THR IAATTTAACTA	ASN THR SER ALA LYS THR VAL THR PHE CAATACTAGCGCTAAAACAGTAACTTTT 1930 ASN ASN VAL L ASN ASN VAL L AACAATGTTA	GLY HIS LYS VAL THR LEU ASN SER LYS CGGTCACACTAAATAGCAAA 1990 VAL LYS THR LI GTGAAACAC	GLU GLY GLY SER ASP ASN ASN THR GLY TGAAGGTGGCAGTGACAACAATACCGGT

	177/235		
A A A 2100	THR C A C 2160	GLY A G G 2220	ASP 2 G A 2280
VAL TA	祖 ここで 2	17HR 1 C A (27.	GLY G C
SLU A A G	ILE THR THR TTACCAC 2160	HIS SATA	ILY G C G
L (IX G A	ALA I	A C
3P VA 1 T G T 2090	JY GI 3 A G G 2150	R AI C C G C 2210	AL AL P G G C 2270
ASP G A 3	GLY G G 7	THR A C (VAL G T (
LYS A A A	ASN A A T	ILE THR ATAACC 2210	THR ILE VAL ALA GLY GLY CAATTGTGGCAGGCG 2270
THR ALA LYS ASP VAL GLU VAL CTGCAAAGATGTAGAAGTA? 2080 2090 2	ALA G C A	THR A C C ASN VAL GLU A C G T G G A G A 2200	THR A C A
THR A C T 2080	 ALA G C G 2140	 VAL G T G 2200	 VAL G T G
ILE A T T A	VAL G T G SER A	THR. ACC. ASN ASN AB AC	SER LYS CCAAG GLY SER V GCTCTG
. LEU THR ILE THR ALA LYS ASP VAL GLU VAL AST . TTAACTATTACTGCAAAAGATGTAGAAGTAAA .2070 2080 2100	ASN ASN ILE THR SER HIS LYS THR VAL CAACAATATTACTTCACAAAACAGTG 2110 ASN VAL SER ALA ALA ASN GLY GLY ILE THR THE AACGTCTCTGCGCAAATGGAGGGATTACCAC AACGTCTCTGCGCCAAATGGAGGGATTACCAC AACGTCTCTGCGCCAAATGGAGGGATTACCAC	LYS THR GLY THR THR ILE ASN ALA THR TAAAACAGGTACAACCATTAATGCAACC A180 ALA GLY ASN VAL GLU ILE THR ALA HIS THR GLY GCCGGTAACGTGAAGATAACGCGTTCATACAGG GCCGGTAACGTGAAGATAACGTCATACAGA	
LEU T T A	S LYS THR C A A A A C A 2120 ASN VAL A A C G T C	E ASN T A A T 2180 ALA G C C	E GLU T G A G 2240 . PRO . C C T . 2250
	HIS C A C 212 	ILE AS A T T A A 2180 AL G C	ILE (A T G 2240 F C C
	THR SER HIS ACTICTCACA 2120 P	THR THR	GLY G G A
	THR ACT	THR A C A	ILE GIN GLY GLY ILE GLU TCCAAGGCGGAATTGAGT 2230 2240 PRO CCTG
	ILE A T T A 2110	GLY 3 G T 7 2170	GLN 2 A A 2230
	ASN A A T A	LYS THR AAACAO	ILE 1 T C (
	ASN A C A	YS A A A	SER . G T A
	C A	T A	C A

SER A G 2460

⊱

ACT

AAAGCA

AACGTT

ACAGTA

LYS

GLY

... G G C A A G *P*

ACCACCCAAGCAGACTCAAAATT... 2470 2480 ...

TTA

⊱

SE

GLN

ASIN AA

ALA

VAL

ALA 1 G C 2340	GLN C A 2400	
GLY ; G T (SER AGT 2	
SER GCO	SER	
ASN A T A	景 CII	
MA AS CAA 2330	所 正 CCAC 2390	
HR P	LE 1 TAA	
L T	R I A I	
VA L G T	SE SE	
ASN ILE SER GLY A T A T T T C A G G C 2300 ASN ALA VAL THR VAL THR ALA ASN SER GLY ALA A A C G C C G T T A C T G C A A A T A G C G G T G C C C C C C C C C C C C C C	GLY SER THR ILE GCTCTACAATT 2360 LYS GLY THR GLU SER ILE THR THR SER GLN AAAGGAACCGAGAGTATAACCACTTCAAGTCA 2380 2400	GLY LYS ILE SER G T A A A A T T T C C 2420
3 A G C G C G C G C G C G C G C G C G C G	R I AA Y TI AA	E L
SEI TC ALA	THI LAC GL)	ILI AAT
AT A T T Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z	LY SER G C T C T 2360 LYS A A A	LYS A A A 20
ASN A A T 23(GLY G G C 23(GLY L G G T A . 2420
GLY G G T	ALA G C A	G G C
VAL G T A	LEU TTGG	ILE A T C (
ALA G C T 2290	THR A C T 2350	SER GLY ASN CAGGTAATA 2410
CTTC	LEU THR TAACC	G G T
THR LEU ALA VAL GLY ASN ILE SER GLY TACTCTTGCTGTAGGTAATATTCAGGC 2290 ASN ALA VAL T ASN ALA VAL T A C C C C C T A C	LEU THR IEU ALA GLY SER THR ILE A T T A A C C A C T T T G G C A G G C T C T A C A A T T 2350 2350 LYS GLY THR G A A A G G A A C C G S 2370	SER GLY ASN ILE GLY GLY LYS ILE SER A T C A G G T A A A T T T C C 2420
	✓ UBSTITUTE SHEET (RU)	
3	ひとりょうしん しょうしんしゅんしゅ	,

178/235

7	=
1	_
C	Ż
(5
ī	Ī

	179 / 235		
A SER A A G 2520	GLU C G A 2580	SER A A G 2640	LYS 2 A A 2700
THR ACA 2	THR	SER TCAP	THR ACC 2
VAL GTA?	ALA G C G	THR A C T	ALA 3 C A ,
ALA ASN CTAATC 2510	VAL THR ALA THR GLU 3 T T A C C G C G A C C G A 2570 2580	THR A C C.	SER SER
ALA AGO TO AGO T	VAL 1 3 T T A 2570	VAL 7 3 T G A 2630	VAL SER ALA THR FTAGCGCAACCA 27
GLU 3 A G (GLU 3 A A C	SER VAL THR THR 1 G C G T G A C C A C T T 2630	GLU AAG
GLY 3 G C (VAL	GLU ; A G Þ	VAL T A G
GLU ALA THR GLU GLY GLU ALA ASN VAL THR SEI G A A G C G A G G C G A G G C T A A T G T A A C A A G G C G A G G C T A A T G T A T A	LYS THR SER ILE ILE GLY GLY THR ILE CAAAACAAGCATAATTGGCGGTACAATT 2530 SER GLY GLY THR VAL GLU VAL THR ALA THR GLA SER GLY GLY THR VAL GLU VAL THR ALA THR GLA TCTGGTGCACAGTAGAAGTTACCGCGACCGA 2580	GLY LEU THR THR GLN ALA GLY SER THR AGGTTTAACCACCCAAGCAGGCTCTACG 2590 ILE THR GLY THR GLU SER VAL THR THR SER SET ATTACTGGAACCGAGGGTGACCACTTCAAG ATTACTGGAACCGAGGGTGACCACTTCAAG	GIN SER GLY ASN ILE GLY GLY MET ILE CCAATCAGGTAATATCGGCGGCATGATT 2650 SER GLY GLY LYS VAL GLU VAL SER ALA THR LY TCTGGTGGCAAAGTAGAAGTTAGCGCAACCAA 2700
	LYS THR AAAACAA	GLY LEU GGTTTA?	GIN SER GLY CAATCAGGTA 2650
	O	A.	Ü

開

MAL

A A C 2760

					_	 CT	.~.	TE :	cur	:==	/61	II E	26			
Ä	A G					ഗ	A A						E ⊶i	TA		
ASP	AT					段	GT (捝	CT(
	TTA					ALA	GCA						THR GLY	C C C		
HI	AGATTTAATT	2710				開	AAGTGCAACA	2770					ASP	TACTGGCGAT	2830	
							1 G G 1						国			
Ϊ	TAA					呂	TAC						国	A A C		
ഗ	AТ					H W	AA						γ γ	J E		
題	000	2720	:	A	2730	田	TTG	2780	:	A	2790		AL.	TTG	2840	
GLY	G T	_	II.E		30	ASP	AC		IFE	L L	06		GLU	AA	,	
Ŕ	TCA		LYS	AAA		GLY	GGT		SER	TCC			ASP	GAT		
THR LYS SER GLY SER GLU	ACTAAATCCGGTTCAGAG		ALA	GCA		GLY THR ILE ASP GLY THR	GGTACAATTGACGGTACG		GLY	GGT			LEU THR VAL GLU ASP ALA	TTAACTGTTGAAGATGCC		
:	:	:	開	ACG	2740	:	:	:	ASIN	AAT	2800		:	:	:	
			ALA	005					閨	ACG						
			GLY	3 G G (VAL	GT						
			ILE LYS ALA THR ALA GLY GLU VAL ASN	AȚTAAAGCAACGGCGGCGAGGTGAATG	27				ILE SER GLY ASN THR VAL ASN VAL THR	ATTTCCGGTAATACGGTAAATGTTACAG	28					
			VAL	G T	2750				VAL	G T	2810					
			ASIN	SAA					選	r A C						
				TG					- 4	A G						

180/235

ASIN

. A A 2820

ø

ე ე

3 A C 2880 G G AGGA(GLY GLY ෆ AC GAT ... G C A *I* ...2850 題 盟

A C C... TTA G G C A A A 7 2900 T C G ACA G C A . 2890 ACT Ø, E $^{\rm L}$ \mathcal{O}

THR LYS ALA SER SER SER ILE THR SER ALA ASN A C T A A G G C T A G T T C A A C C A T T A C T T C A G C T A A 2920 2930 2940	ASN GIN VAL ASN LEU SER ALA LYS ASP TAACCAGGTAAACCTTTCAGCTAAGGAT 2950 GLY SER ILE GLY GLY ASN ILE ASN ALA ALA ASN 88 GGTAGCATTGGGGGAAATATCAATGCTGCTAAA 87 2950 2950 2950 2950	VAL THR LEU ASN THR THR GLY ALA LEU TGTAACACTGAATACTACAGGCGCTCTA 3020 THR THR VAL LYS GLY SER SER ILE ASN ALA ASN ACTACCGTGAAGGGTTCAAGCATTAACGCAAA 3050 3060	SER GLY THR LEU VAL·ILE ASN ALA LYS CAGCGCACCTTGGTTATTAACGCAAAA 3070 ASP ALA GLU LEU ASN GLY GLU ALA SER GLY ASN GACGCTGAGCTAAATGGTGAGCATCAGGTAA 3120
	ASN	VAL	SER
	TAAC	TGTA	CAGC

	丁
1	_
	V.
(j
ī	工

	182/235		
HIS THR VAL VAL ASN ALA THR ASN ALA CCATACAGTAGTGAATGCAACCCA 3130 ASN GLY SER GLY SER VAL ILE ALA THR THR SER ASN GLY SER GLY SER VAL ILE ALA THR THR 3180 AATGGCTCCGGCAGCGTAATCGCGACAACCTC 3180	SER ARG VAL ASN ILLE THR GLY ASP LEU A A G C A G A G T G A A C A T C A C T G G G A T T T A 3190 ILLE THR ILLE ASN GLY LEU ASN ILLE ILLE SER LYS A T C A C A A T A A A T G G A T T A A A T A C A T T T C A A A A A A	ASN GLY ILE ASN THR VAL LEU LEU LYS A A A C G G T A C T G T T A A A A 3250 GLY VAL LYS ILE ASP VAL LYS TYR ILE GIN PRO G C C T T A A A T T G A A T A C A T T C A A C C 3280 3290 3300	GLY ILE ALA SER VAL ASP GLU VAL ILE GGGTATAGCAAGCGTAGAAGTAATT 3310
HIS CATA	SER A G C A	ASN (GLY]
O	K	A	Ŋ

FIG.27Q

JYS ASP A A G A 3360	HE ALA SHE T T G C 3420	SER CAAG 3480	LE CYS TATG 3540
GLU ALA LYS ARG ILE LEU GLU LYS VAL LYS ASI G A A G C G A A A C G C A T C C T T G A G A G G T A A A G A 3340 3350 3360	LEU SER ASP GLU GLU ARG GLU ALA LEU T T T A T C T G A T G A A G A G A G C G T T A 3370 ALA LYS LEU GLY VAL SER ALA VAL ARG PHE ALA G C T A A A C T T G C G T A A G C G T T T T T G C 3420	GLU PRO ASN ALA ILE THR ILE ASN TGAGCCAATTACGATTAAT 3430 THR GLN ASN GLU PHE THR THR ARG PRO SER SES ACACAAAATGAGTTTACAACCAGACCATCAAG 3480 3480	GIN VAL THR ILE SER GLU GLY LYS VAL TCAAGTGACAATTTCTGAAGGTAAGGTA 3490 CYS PHE LEU ILE GLY ASN GLY ALA THR ILE CYS TGTTTCTTAATCGGCAATGGTGCAACAATATG 3530 3540
	LEU TTTAT	GLU] TGAGC	GIN TCAAG

184/235

FIG.Z/H
THR ASN ILE ALA ASP ILE GLU ARG
CACCAATATTGCTGATATTGAGCGG
3550
3550

60 120	180	300	404	452
сасститите стастетата пставаратать прададата стагарадис сссататада	ARGSTATAAT CITTCATCIT TCATCITTCA TCITTCATCT TTCATCTTTC ATCITTTCATCTTTCATCTTTCATCTTTTCATCTTTTCATCTTTTCATCTTTTCATCTTTTCATCTTTTCATCTTTTCATCTTTTCATCTTTTTCATCTTTTTCATCTTTTTCATCTTTTTCATCTTTTTT	ACATIGOCCTIC ATGAACCGAG GGAAGGGAGG GAGGGGAAG AATGAAGAGG GAGCTGAACG AACGCAAATG ATAAAGTAAT TTAATTGTTC AACTAACCTT AGGAGAAAAT ATG AAC Met Asn 1	AAG CTA TAT CGT CTC AAA TTC AGC AAA CGC CTG AAT GCT TTG GTT GCT Lys Leu Tyr Arg Leu Lys Phe Ser Lys Arg Leu Asn Ala Leu Val Ala 5	GTG TCT GAA TTG GCA CGG GGT TGT GAC CAT TCC ACA GAA AAA GGC AGC Val Ser Glu Leu Ala Arg Gly Cys Asp His Ser Thr Glu Lys Gly Ser 20

FIG.28E

200	548	296	644	692	740
GAA AAA CCT GCT CGC ATG AAA GTG CGT CAC TTA GCG TTA AAG CCA CTT Glu Lys Pro Ala Arg Met Lys Val Arg His Leu Ala Leu Lys Pro Leu 35 40 50	TCC GCT ATG TTA CTA TCT TTA GGT GTA ACA TCT ATT CCA CAA TCT GTT Ser Ala Met Leu Leu Ser Leu Gly Val Thr Ser Ile Pro Gln Ser Val 55 60	TTA GCA AGC GGC TTA CAA GGA ATG GAT GTA GTA CAC GGC ACA GCC ACT Leu Ala Ser Gly Leu Gln Gly Met Asp Val Val His Gly Thr Ala Thr 70 75 80	AIG CAA GTA GAT GGT AAT AAA ACC ATT ATC CGC AAC AGT GTT GAC GAT Met Gln Val Asp Gly Asn Lys Thr Ile Ile Arg Asn Ser Val Asp Asp 85	ATC ATT AAT TGG AAA CAA TTT AAC ATC GAC CAA AAT GAA ATG GTG CAG Ile Ile Asn Trp Lys Gln Phe Asn Ile Asp Gln Asn Glu Met Val Gln 100	TITY TITA CAA GAA AAC AAC TCC GCC GITA TITC AAC CGT GITY ACA TCT. Phe Leu Gln Glu Asn Asn Asn Ser Ala Val Phe Asn Arg Val Thr Ser 115
	SUBS	STITUTE SHEE	I (NULE 40)		

788	836	884	932	980	1028			
AAC CAA ATC TCC CAA TTA AAA GGG AFF TTA GAF TCF AAC GGA CAA GTC Asn Gln Ile Ser Gln Leu Lys Gly Ile Leu Asp Ser Asn Gly Gln Val 135	TIT TTA AIC AAC CCA AAT GGF ATC ACA ATA GGF AAA GAC GCA AIT AIT Phe Leu Ile Asn Pro Asn Gly Ile Thr Ile Gly Lys Asp Ala Ile Ile 150	AAC ACT AAT GGC TITT ACG GCT TCT ACG CTA GAC ATT TCT AAC GAA AAC Asn Thr Asn Gly Phe Thr Ala Ser Thr Leu Asp Ile Ser Asn Glu Asn 175	ATC AAG GCG CGT AAT TTC ACC TTC GAG CAA ACC AAA GAT AAA GCG CTC Ile Lys Ala Arg Asn Phe Thr Phe Glu Gln Thr Lys Asp Lys Ala Leu 180	GCT GAA ATT GTG AAT CAC GGT TTA ATT ACT GTC GGT AAA GAC GGC AGT Ala Glu Ile Val Ash His Gly Leu Ile Thr Val Gly Lys Asp Gly Ser 195	GTA AAT CTT AIT GGT GGC AAA GTG AAA AAC GAG GGT GTG AITT AGC GTA Val Asn Leu Ile Gly Gly Lys Val Lys Asn Glu Gly Val Ile Ser Val 220			

188/235

	1316	1268	1220	1172	1124	1076
Ala Glu Ile Gly Gly Val Ile Ser Ala Gln Asn Gln 315	GGC GGT GTA ATT TCC GCT CAA AAT CAG CAA GCT	GTA AGC AAA GAT AAA AGC GGC AAT ATT GTT CTT TCC GCC AAA GAG GGT Val Ser Lys Asp Lys Ser Gly Asn Ile Val Leu Ser Ala Lys Glu Gly 295	GIC Val	GAA GCG GTC AAT CTG GGC GAT ATT TTT GCC AAA GGC GGT AAC ATT AAT Glu Ala Val Asn Leu Gly Asp Ile Phe Ala Lys Gly Gly Asn Ile Asn 260	GAT ATA ATA AAC CCA ACC ATT ACT TAC AGC ATT GCC GCG CCT GAA AAT ASP Ile Ile ASn Pro Thr Ile Thr Tyr Ser Ile Ala Ala Pro Glu Asn 255	AAT GGT GGC AGC AIT ICT TIA CIC GCA GGG CAA AAA AIC ACC AIC AGC Asn Gly Gly Ser Ile Ser Leu Leu Ala Gly Gln Lys Ile Thr Ile Ser 230
Ala Glu Ile Gly Gly Val Ile Ser Ala Gln Asn Gln Gln Ala	GCG GAA ATT GGC GGT GTA ATT TCC GCT CAA AAT CAG CAA GCT	AGC AAA GAT AAA AGC GGC AAT ATT GTT CTT TCC GCC AAA GAG Ser Lys Asp Lys Ser Gly Asn Ile Val Leu Ser Ala Lys Glu 295	GIC CGT GCT GCC ACT ATT CGA AAC CAA GGT AAA CTT TCT GCT GAT Val Arg Ala Ala Thr Ile Arg Asn Gln Gly Lys Leu Ser Ala Asp 275	GAA GCG GTC AAT CTG GGC GAT ATT TTT GCC AAA GGC GGT AAC ATT Glu Ala Val Asn Leu Gly Asp Ile Phe Ala Lys Gly Gly Asn Ile 260	GAT AITA ATA AAC CCA ACC AITT ACT TAC AGC AITT GCC GCG CCT GAA ASP Ile Ile ASN Pro Thr Ile Thr Tyr Ser Ile Ala Ala Pro Glu 245	GGT GGC AGC AIT TOT TIPA CTC GCA GGG CAA AAA ATC ACC ATC Gly Gly Ser Ile Ser Leu Leu Ala Gly Gln Lys Ile Thr Ile 230 235

FIG.28E

1364	1412	1460	1508	1556	1604		
GGC GGC AAG CTG ATG ATT ACA GGC GAT AAA GTC ACA TTA AAA ACA GGT Gly Gly Lys Leu Met Ile Thr Gly Asp Lys Val Thr Leu Lys Thr Gly 325	GCA GIT AIC GAC CIT ICA GGT AAA GAA GGG GGA GAA ACT IAC CIT GGC Ala Val Ile Asp Leu Ser Gly Lys Glu Gly Gly Glu Thr Tyr Leu Gly 340 340	GET GAC GAG CGC GAA GGT AAA AAG GGC ATT CAA TTA GCA AAG AAA Gly Asp Glu Arg Gly Glu Gly Lys Lys Gly Ile Gln Leu Ala Lys Lys 355 360 370	ACC TOT TTA GAA AAA GGO TOA AOO ATO AAT GTA TOA GGO AAA GAA AAA Thr Ser Leu Glu Lys Gly Ser Thr Ile Asn Val Ser Gly Lys Glu Lys 375	GCC GCA CGC GCT ATT GTG TGG GGC GAT ATT GCG TTA ATT GAC GGC AAT Gly Gly Arg Ala Ile Val Trp Gly Asp Ile Ala Leu Ile Asp Gly Asn 390 400	ATT AAC GCT CAA GGT AGT GGT GAT ATC GCT AAA ACC GGT GGT TTT GTG Ile Asn Ala Gln Gly Ser Gly Asp Ile Ala Lys Thr Gly Gly Phe Val 405		
	SUBSTITUTE SHEET (RULE 26)						

FIG.28

1652	1700	1748	1796	1844	1892
	(-1	(-1	(-1		 1
3 GGG CAT GAT TTA TTC ATC AAA GAC AAT GCA ATT GTT GAC 1 Gly His Asp Leu Phe Ile Lys Asp Asn Ala Ile Val Asp 425 1	FIGG TITG TITA GAC COG GAT AAT GITA TOTT AITT AAT GCA GAA 1 Trp Leu Leu Asp Pro Asp Asn Val Ser Ile Asn Ala Glu 440	A CGC AGC AAT ACT TCA GAA GAC GAT GAA TAC ACG GGA TCC / Arg Ser Asn Thr Ser Glu Asp Asp Glu Tyr Thr Gly Ser 455	r GCC AGC ACC CCA AAA CGA AAC AAA GAA AAG ACA ACA	A ACT CTT GAG AGT ATA CTA AAA AAA GGT ACC TTT GTT AAC : Thr Leu Glu Ser Ile Leu Lys Lys Gly Thr Phe Val Asn 5 495	r AAT CAA CGC ATC TAT GTC AAT AGC TCC ATT AAT TTA TCC a Asn Gln Arg Ile Tyr Val Asn Ser Ser Ile Asn Leu Ser 505
ACG TCG Thr Ser 420	AAA GAG Lys Glu	GCA GGA (Ala Gly ,	AAT AGT Asn Ser	AAC ACA Asn Thr 485	ACT GCT Thr Ala 500
GAG A(Glu II	GCC Al Ala Ly 435	ACA G Thr A	GGG AA Gly As	ACA A2 Thr As	ATC AC Ile TP 5(
		STITUTE SHE			

FIG.28(

1940	1988	2036	2084	2132	2180
AAT GGC AGC TTA ACT CTT TGG AGT GAG GGT CGG AGC GGT GGC GGT GT ASN Gly Ser Leu Thr Leu Trp Ser Glu Gly Arg Ser Gly Gly Gly Val 515 520	GAG ATT AAC AAC GAT ATT ACC ACC GGT GAT GAT ACC AGA GGT GCA AAC	TTA ACA ATT TAC TCA GGC GGC TGG GTT GAT GTT CAT AAA AAT ATC TCA	CTC GGG GCG CAA GGT AAC ATA AAC ATT ACA GCT AAA CAA GAT ATC GCC	TIT GAG AAA GGA AGC AAC CAA GIC ATT ACA GGT CAA GGG ACT ATT ACC	TCA GGC AAT CAA AAA GGT TTT AGA TTT AAT AAT GTC TCT CTA AAC GGC
	Glu Ile Asn Asp Ile Thr Thr Gly Asp Asp Thr Arg Gly Ala Asn	Leu Thr Ile Tyr Ser Gly Gly Trp Val Asp Val His Lys Asn Ile Ser	Leu Gly Ala Gln Gly Asn Ile Asn Ile Thr Ala Lys Gln Asp Ile Ala	Phe Glu Lys Gly Ser Asn Gln Val Ile Thr Gly Gln Gly Thr Ile Thr	Ser Gly Asn Gln Lys Gly Phe Arg Phe Asn Asn Val Ser Leu Asn Gly
	535 545	550 555	570 575	580 585	595

SUBSTITUTE SHEET (RULE 26)

IG.28H

2228	2276	2324	2372	2420	2468
ACT GGC AGC GGA CTG CAA TTC ACC ACT AAA AGA ACC AAT AAA TAC GCT Thr Gly Ser Gly Leu Gln Phe Thr Thr Lys Arg Thr Asn Lys Tyr Ala 625	ATC ACA AAT AAA TTT GAA GGG ACT TTA AAT ATT TCA GGG AAA GTG AAC Ile Thr Asn Lys Phe Glu Gly Thr Leu Asn Ile Ser Gly Lys Val Asn 630 630	ATC TCA ATG GTT TTA CCT AAA AAT GAA AGT GGA TAT GAT AAA TTC AAA Ile Ser Met Val Leu Pro Lys Asn Glu Ser Gly Tyr Asp Lys Phe Lys 645	GGA CGC ACT TAC TGG AAT TTA ACC TCC TTA AAT GTT TCC GAG AGT GGC Gly Arg Thr Tyr Trp Asn Leu Thr Ser Leu Asn Val Ser Glu Ser Gly 660 665	GAG TITT AAC CTC ACT ATT GAC TCC AGA GGA AGC GAT AGT GCA GGC ACA Glu Phe Asn Leu Thr Ile Asp Ser Arg Gly Ser Asp Ser Ala Gly Thr 675 689	CIT ACC CAG CCT TAT AAT TTA AAC GGT ATA TCA TTC AAC AAA GAC ACT Leu Thr Gln Pro Tyr Asn Leu Asn Gly Ile Ser Phe Asn Lys Asp Thr 695

192/235

FIG.28

2516 2564 2612 2660 2708 2756 CTT CTC Leu Leu Asn 8 AAA Lyss 770 G1yAAT GCC AAG Lys Phe Ser TCA Ser 785 Asn GIT ATA AAT T TITT AAA ACT The Phe Lys Thr Ile 720 ACA Thr TTA Ser Leu 800 GAC Asp IIC Phe Ala 735 TTA ACT Thr Tyr, Phe Asp 750 TIT GAT TAC Ferr Asn Arg AAC AAT Asn GGG GGG AGT GTT Gly Gly Ser Val AGA GIA Val 765 Asp GAG AAA GAT Lys. Val TIG TTA Leu 780 GGT Gly Leu CAA ACC CCC (Glu Thr Pro (GGG TCA AGT Coly Ser Ser 1 AGA Arg 715 Ser AGT Glu 795 Thr Ala CGA AAT GCA ATA Ser 730 GGA G1y 745 Ser. Glu Arg Asn TyrTAT <u>1</u> Ser ACA Thr GIC ATA AAT AAG Ile Asn Lys TIC Phe Val 760 Val GAA GIT Asn TCA Ser 775 AAC 9 Gly Ser Ser Val TCA GIJ Asn Val 710 Thr 790 AAT Asn 999 Ile Gly 725 AAA Lys Ser TIT AAT TIL AAC Asn Phe lle ATA TCA Ser Phe 740 Pro SEA Ala 755

SUBSTITUTE SHEET (RULE 26)

FIG.28

2804	2852	2900	2948	2996	3044
ACC GGA GGC AAC ATA ACA CTT TTG CAA GTT GAA GGC ACC GAT GGA ATG	ATT GGT AAA GGC ATT GTA GCC AAA AAA AAC ATA ACC TTT GAA GGA GGT	AAC ATC ACC TTT GGC TCC AGG AAA GCC GTA ACA GAA ATC GAA GGC AAT	GIT ACT AIC AAT AAC AAC GCT AAC GIC ACT CIT AIC GGT ICG GAT TITT	GAC AAC CAT CAA AAA CCT TTA ACT ATT AAA AAA	AGC GGC AAC CTT ACC GCT GGA GGC AAT ATT GTC AAT ATA GCC GGA AAT
Thr Gly Gly Asn Ile Thr Leu Leu Gln Val Glu Gly Thr Asp Gly Met	Ile Gly Lys Gly Ile Val Ala Lys Lys Asn Ile Thr Phe Glu Gly Gly	Asn 11e Thr Phe Gly Ser Arg Lys Ala Val Thr Glu 11e Glu Gly Asn	Val Thr Ile Asn Asn Asn Ala Asn Val Thr Leu Ile Gly Ser Asp Phe		Ser Gly Asn Leu Thr Ala Gly Gly Asn Ile Val Asn Ile Ala Gly Asn
805	820	835	855		885

SUBSTITUTE SHEET (RULE 26)

FIG.28k

3092	3140	3188	3236	3284	3332
CIT ACC GIT GAA AGT AAC GCT AAT TIC AAA GCT AIC ACA AAT TIC ACT Leu Thr Val Glu Ser Asn Ala Asn Phe Lys Ala Ile Thr Asn Phe Thr 900	TIT AAT GTA GGC GGC TTG TTT GAC AAC AAA GGC AAT TCA AAT ATT TCC Phe Asn Val Gly Leu Phe Asp Asn Lys Gly Asn Ser Asn Ile Ser 915	ATT GCC AAA GGA GCT CGC TTT AAA GAC ATT GAT AAT TCC AAG AAT Ile Ala Lys Gly Ala Arg Phe Lys Asp Ile Asp Asn Ser Lys Asn 935	TITA AGC ATC ACC AAC TCC AGC TCC ACT TAC CGC ACT ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC ACT TAC CGC ACT ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC ACT TAC CGC ACT ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC ACT TAC CGC ACT ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC AGC TCC ACT TAC ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC AGC TCC ACT TAC OCT ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC AGC TCC ACT TAC OCT ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC AGC TCC ACT TAC OCT ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC AGC TCC ACT TAC OCT ATT ATA AGC 1. TTA AGC ATC ACC AAC TCC AGC TCC ACT TAC OCT ACT ATT ATA AGC 1. TTA AGC ATC ACC ACC AAC TCC AGC TCC ACT TAC OCT ACT ATT ATA AGC 1. TTA AGC ATC ACC ACC AAC TCC AGC TCC ACT TAC OCT ACT ATT ATA AGC 1. TTA AGC ATC ACC ACC AAC TCC AGC TCC ACT TAC OCT ACT ATT ATA AGC 1. TTA AGC ATC ACC ACC AAC TCC AGC TCC ACT TAC OCT ACT ATT ATA AGC 1. TTA AGC ACT ACT ATT ATT ATT AGC 1. TTA AGC ACC ACC ACC AAC TCC AGC TCC ACT TAC OCT ACT ATT ATT ATT AGC 1. TTA AGC ACT ACT ACT ACT ATT ATT ATT ATT AGC 1. TTA AGC ACT ACT ACT ACT ATT ATT ATT ATT ATT AT	GGC AAT ATA ACC AAT AAA AAC GGT GAT TTA AAT AIT ACG AAC GAA GGT Gly Asn Ile Thr Asn Lys Asn Gly Asp Leu Asn Ile Thr Asn Glu Gly 975	AGT GAT ACT GAA ATG CAA ATT GGC GGC GAT GTC TCG CAA AAA GAA GGT 3. Ser Asp Thr Glu Met Gln Ile Gly Gly Asp Val Ser Gln Lys Glu Gly 980

FIG.28L

S AAA ATC AAT ATT ACC AAA CAG A Lys Ile Asn Ile Thr Lys Gln 1005	dia Asn Ser Asp Ser As 1020	AAA ACC AAA GAA TIG AAA TIA ACG CAA GAC Lys Thr Lys Glu Leu Lys Leu Thr Gln Asp 1035	AAT AAA GCA GAG ATT ACA GCT AAA GAT GGT Asn Lys Ala Glu Ile Thr Ala Lys Asp Gly 1050	AAC ACC AAT AGT GCT GAT GGT ACT AAT Asn Thr Asn Ser Ala Asp Gly Thr Asn 1065	TIT AAC CAG GIT AAA GAT TCA AAA ATC TCT GCT GAC Phe Asn Gln Val Lys Asp Ser Lys Ile Ser Ala Asp 1080
	All AAG GLA GGI GAI GAS GAS AAG ICC GAI ICA GAC GCG Ille Lys Ala Gly Val Asp Gly Glu Asn Ser Asp Ser Asp Ala '1015	AAT GCC AAT CTA ACC AIT AAA Asn Ala Asn Leu Thr Ile Lys 1030	CTA AAT ATT TCA GGT TTC AAT Leu Asn Ile Ser Gly Phe Asn 1045	AGT GAT TTA ACT ATT GGT AAC I Ser Asp Leu Thr Ile Gly Asn 1065	AAA AAA GTA ACC TTT AAC CAG Lys Lys Val Thr Phe Asn Gln 1075
AAT Asn 995		AAT ASIN ASIN	CE PE	AGT	AAA Lys 1075

FIG.28M

	₩ Ħ				
CAC AAG GTG ACA CTA CAC AGC AAA GTG GAA ACA TCC GGT AGT AAT His Lys Val Thr Leu His Ser Lys Val Glu Thr Ser Gly Ser Asn 1095	AAC AAC ACT GAA GAT AGC AGT GAC AAT AAT GCC GGC TIA ACT ATC GAT Asn Asn Thr Glu Asp Ser Ser Asp Asn Asn Ala Gly Leu Thr Ile Asp 1110	GCA AAA AAT GTA ACA GTA AAC AAC AAT ATT ACT TCT CAC AAA GCA GTG Ala Lys Asn Val Thr Val Asn Asn Ile Thr Ser His Lys Ala Val 1125	AGC ATC TCT GCG ACA AGT GGA GAA ATT ACC ACT AAA ACA GGT ACA ACC Ser Ile Ser Ala Thr Ser Gly Glu Ile Thr Thr Lys Thr Gly Thr Thr 1140	ATT AAC GCA ACC ACT GGT AAC GTG GAG ATA ACC GCT CAA ACA GGT AGT Ile Asn Ala Thr Thr Gly Asn Val Glu Ile Thr Ala Gln Thr Gly Ser 1155	ATC CTA GGT GGA ATT GAG TCC AGC TCT GGC TCT GTA ACA CTT ACT GCA Ile Leu Gly Gly Ile Glu Ser Ser Ser Gly Ser Val Thr Leu Thr Ala 1185
3668	3716	3764	3812	3860	3908

FIG.28N

3956 4052 4100 4148 4196 1250 Lys Gly Thr Glu Ser Val Thr Thr Ser Ser Gln Ser Gly Asp Ile Gly 1220 3GT ACG ATT TCT GGT GGC ACA GTA GAG GTT AAA GCA ACC GAA AGT TTA 3ly Thr Ile Ser Gly Gly Thr Val Glu Val Lys Ala Thr Glu Ser Leu 큠 CAA TCC AAT TCA AAA ATT AAA GCA ACA ACA GGC GAG GCT AAC Thr Gln Ser Asn Ser Lys Ile Lys Ala Thr Thr Gly Glu Ala Asn 1255 GTA ACA AGT GCA ACA GGT ACA ATT GGT GGT ACG ATT TCC GGT AAT ACG Val Thr Ser Ala Thr Gly Thr Ile Gly Gly Thr Ile Ser Gly Asn Thr GGC AAC ACC GIT ACT GGT GCA TITA ACC ACT TITG GCA GGC TCT ACA AIT GAG AGT GITA ACC ACT TCA AGT CAA TCA GGC GAT ATC GGC Thr Gly Asn Thr Val Gly Ala Leu Thr Thr Leu Ala Gly Ser 1210 <u>1</u>00 Ser Asn Ile Ser ACC GAG GGC GCT CTT GCT GTA AGC AAT ATT Ala Leu Ala Val ACT GCA AAT AGC Ala Asn Ser Glu Gly AAA GGA ACC ACT Thr Gly Thr ACC. Val SUBSTITUTE SHEET (RULE 26)

FIG.28(

FIG.28F

4532	4580	4628	4676	4724	4772
GCT GAG CTA AAT GGC GCA GCA TTG GGT AAC CAC ACA GTG GTA AAT GCA Ala Glu Leu Asn Gly Ala Ala Leu Gly Asn His Thr Val Val Asn Ala 1380	ACC AAC GCA AAT GGC TCC GGC AGC GTA ATC GCG ACA ACC TCA AGC AGA Thr Asn Ala Asn Gly Ser Gly Ser Val Ile Ala Thr Thr Ser Ser Arg 1395	GIG AAC ATC ACT GGG GAT TTA ATC ACA ATA AAT GGA TTA AAT ATC ATT Wal Asn Ile Thr Gly Asp Leu Ile Thr Ile Asn Gly Leu Asn Ile Ile 1415	TCA AAA AAC GGT ATA AAC ACC GTA CTG TTA AAA GGC GTT AAA ATT GAT F Ser Lys Asn Gly Ile Asn Thr Val Leu Leu Lys Gly Val Lys Ile Asp 1430 1430	GIG AAA TAC ATT CAA CCG GGT ATA GCA AGC GTA GAT GAA GTA ATT GAA Val Lys Tyr Ile Gln Pro Gly Ile Ala Ser Val Asp Glu Val Ile Glu 1455	GCG AAA CGC ATC CTT GAG AAG GTA AAA GAT TTA TCT GAT GAA GAA AGA Ala Lys Arg Ile Leu Glu Lys Val Lys Asp Leu Ser Asp Glu Glu Arg 1460

5116

AGAATACAAT AAAGTATTTT TAACAGGITTA TTATTATG

FIG.28Q

4820 4868 4916 5078 4958 5018 TTATTTACTG TGTGGGTTAA AGITCAGTAC GGGCTTTACC CATCTTGTAA AAAATTACGG TAGCOGICAG TAATTGACAA GGTAGATTTC ATCCTGCAAT GAAGICATTT TATTTTCGTA 1490 Arg Pro 1505 GAG CCA Glu Pro AAT AAT ACA AIT ACA GIC GAT ACA CAA AAT GAA TIT GCA ACC AGA CCA AGT CGA ATA GTG ATT TCT GAA GGC AGG GCG TGT TTC TCA AAC AGT Ile Val Ile Ser Glu Gly Arg Ala Cys Phe Ser Asn Ser 1510 1520 Ile (Thr GAT GCC GCC ACC GTG TGC GTT AAT ATC GCT GAT AAC GGG CGG Gly Ala Thr Val Cys Val Asn Ile Ala Asp Asn Gly Arg Glu Ala Leu Ala Lys Leu Gly Val Ser Ala Val Arg Phe AAA CITI GGA GTA AGT GCT GTA CGT TITI Asp Thr Gln Asn Glu Phe Ala 1500 1480 Ile Thr Val GAA GCG TTIA GCT Ser Arg Asn Asn Thr <u>a</u> SUBSTITUTE SHEET (RULE 26)

FIG.29

			SUB	STITU	TE SHEET	(RULE 26)		
TAAATATACA AGATAATAAA AATAAATCAA GATTITITIGIG AIGACAAACA ACAATTACAA	CACCITITITY GCAGICTATA IGCAAATATT ITAAAAAAAT AGTATAAAIC CGCCATATAA	AAIGGIAIAA ICITICAICI IICAICIIIA AICIIICAIC IIICAICIII CAICIIIICAI	CTITICATOTI TOTATION TOTATION TACATOTITIC ALCHARATO TATOATOTIT	CACATGAAAT GATGAACCGA GGGAAGGGAG GCAGGGGCAA GAATGAAGAG GGAGCTGAAC	GAACGCAAAT GATAAAGTAA TITAATIGIT CAACTAACCT TAGGAGAAAA T ATG AAC Met Asn 1	AAG ATA TAT CGT CTC AAA TTC AGC AAA CGC CTG AAT GCT TTG GTT GCT Lys 11e Tyr Arg Leu Lys Phe Ser Lys Arg Leu Asn Ala Leu Val Ala 5	GTG TCT GAA TTG GCA CGG GGT TGT GAC CAT TCC ACA GAA AAA GGC TTC Val Ser Glu Leu Ala Arg Gly Cys Asp His Ser Thr Glu Lys Gly Phe 20	CGC TAT GIT ACT ATC TIT AGG TGT AAC CAC TTA GCG TTA AAG CCA CITT Arg Tyr Val Thr Ile Phe Arg Cys Asn His Leu Ala Leu Lys Pro Leu As
09	120	180	240	300	357	405	453	501

549	597	645	693	741	789	837
TCC GCT AIG TTA CTA TCT TTA GGT GTA ACA TCT AIT CCA CAA TCT GTT Ser Ala Met Leu Leu Ser Leu Gly Val Thr Ser Ile Pro Gln Ser Val 60	TTA GCA AGC GGC TTA CAA GGA ATG GAT GTA GTA CAC GGC ACA GCC ACT Leu Ala Ser Gly Leu Gln Gly Met Asp Val Val His Gly Thr Ala Thr 70	ATG CAA GTA GAT GGT AAT AAA ACC ATT ATC CGC AAC AGT GTT GAC GCT Met Gln Val Asp Gly Asn Lys Thr Ile Ile Arg Asn Ser Val Asp Ala 85	ATC ATT AAT TGG AAA CAA TTT AAC ATC GAC CAA AAT GAA ATG GTG CAG Ile Ile Asn Trp Lys Gln Phe Asn Ile Asp Gln Asn Glu Met Val Gln 100	IT TIA CAA GAA AAC AAC TCC GCC GI'A TI'C AAC CGT GIT ACA TCT he Leu Gln Glu Asn Asn Ser Ala Val Phe Asn Arg Val Thr Ser 120	AAC CAA ATC TCC CAA TTA AAA GGG ATT TTA GAT TCT AAC GGA CAA GTC Asn Gln Ile Ser Gln Leu Lys Gly Ile Leu Asp Ser Asn Gly Gln Val 135	
	5	SUBSTITUT	E SHEET (R	IULE 26)		

205/235

1221	1269	1317	1365	1413	1461	1509
GTC CGT GCT GCC ACT AIT CGA AAC CAA GGT AAA CTT TCT GCT GAT TCT	GTA AGC AAA GAT AAA AGC GGC AAT ATT GTT CTT TCC GCC AAA GAG GGT	gaa eeg gaa att eec eer eta att tee eer eaa aar eag eaa eer aaa	GCC GCC AAG CTG ATG ATT ACA GCC GAT AAA GTC ACA TTA AAA ACA GGT	GCA GIT AIC GAC CIT ICA GGF AAA GAA GGG GGA GAA ACT IAC CIT GGC	GGT GAC GAG CGC GAC GAA GGT AAA AAC GGC ATT CAA TTA GCA AAG AAA	ACC TCT TTA GAA AAA GGC TCA ACC ATC AAT GTA TCA GGC AAA GAA AAA
Val Arg Ala Ala Thr Ile Arg Asn Gln Gly Lys Leu Ser Ala Asp Ser	Val Ser Lys Asp Lys Ser Gly Asn Ile Val Leu Ser Ala Lys Glu Gly	Glu ala elu 11e ely ely Val 11e Ser ala eln Asn eln eln ala Lys	Gly Gly Lys Leu Met Ile Thr Gly Asp Lys Val Thr Leu Lys Thr Gly	Ala Val Ile Asp Leu Ser Gly Lys Glu Gly Gly Glu Thr Tyr Leu Gly	Gly Asp Glu Arg Gly Glu Gly Lys Asn Gly Ile Gln Leu Ala Lys Lys	Thr Ser Leu Glu Lys Gly Ser Thr Ile Asn Val Ser Gly Lys Glu Lys
275	295	310	325	340	355	375

WO 00/20609

206/235

1557	1605	1653	1701	1749	1797	1845
GGC GGA CGC GCT AITT GTG TGG GGC GAT ATT GCG TTA AITT GAC GGC AAT Gly Gly Arg Ala Ile Val Trp Gly Asp Ile Ala Leu Ile Asp Gly Asn 390	AIT AAC GCT CAA GGT AGT GGT GAT ATC GCT AAA ACC GGT GGT TITT GTG	GAG ACA TCG GGG CAT TAT TTA TCC ATT GAC AGC AAT GCA ATT GTT AAA	ACA AAA GAG TOG TTG CTA GAC CCT GAT GAT GTA ACA ATT GAA GCC GAA	GAC CCC CTT CCC AAT AAT ACC GGT ATA AAT GAT GAA TTC CCA ACA GCC	ACC GGT GAA GCA AGC GAC CCT AAA AAA AAT AGC GAA CTC AAA ACA ACG	CTA ACC AAT ACA ACT ATT TCA AAT TAT CTG AAA AAC GCC TGG ACA ATG
	Ile Asn Ala Gln Gly Ser Gly Asp Ile Ala Lys Thr Gly Gly Phe Val	Glu Thr Ser Gly His Tyr Leu Ser Ile Asp Ser Asn Ala Ile Val Lys	Thr Lys Glu Trp Leu Leu Asp Pro Asp Asp Val Thr Ile Glu Ala Glu	Asp Pro Leu Arg Asn Asn Thr Gly Ile Asn Asp Glu Phe Pro Thr Gly	Thr Gly Glu Ala Ser Asp Pro Lys Lys Asn Ser Glu Leu Lys Thr Thr	Leu Thr Asn Thr Thr Ile Ser Asn Tyr Leu Lys Asn Ala Trp Thr Met
	405	420	435	455	470	485

700 C

1893	1941	1989	2037	2085	2133	2181
AAT ATA ACG GCA TCA AGA AAA CTT ACC GIT AAF AGC TCA ATC AAC ATC Asn Ile Thr Ala Ser Arg Lys Leu Thr Val Asn Ser Ser Ile Asn Ile 500	GGA AGC AAC TCC CAC TTA AIT CTC CAT AGT AAA GGT CAG CGT GGC GGA Gly Ser Asn Ser His Leu Ile Leu His Ser Lys Gly Gln Arg Gly Gly 515	GGC GIT CAG AIT GAT GGA GAT AIT ACT TCT AAA GGC GGA AAT TTA ACC Gly Val Gln Ile Asp Gly Asp Ile Thr Ser Lys Gly Gly Asn Leu Thr 540 545	AIT TAT TCT CCC CCA TCC CIT CAT GIT CAT AAA AAT ATT ACC CIT CAT Ile Tyr Ser Gly Gly Trp Val Asp Val His Lys Asn Ile Thr Leu Asp 550	CAG GGT TITT TITA AAT AITT ACC GCC GCT TCC GTA GCT TITT GAA GGT GGA Gln Gly Phe Leu Asn Ile Thr Ala Ala Ser Val Ala Phe Glu Gly Gly 575	AAT AAC AAA GCA CGC GCG GCA AAT GCT AAA ATT GTC GCC CAG GGC Asn Asn Lys Ala Arg Asp Ala Ala Asn Ala Lys Ile Val Ala Gln Gly 580	ACT GTA ACC ATT ACA GGA GAG GGA AAA GAT TTC AGG GCT AAC AAC GTA Thr Val Thr Ile Thr Gly Glu Gly Lys Asp Phe Arg Ala Asn Asn Val 595

208/235

2229	2277	2325	2373	2421	2469	2517
TCT TTA AAC GGA ACG GGT AAA GGT CTG AAT ATC ATT TCA TCA GTG AAT Ser Leu Asn Gly Thr Gly Lys Gly Leu Asn Ile Ile Ser Ser Val Asn 620	AAT TTA ACC CAC AAT CTT AGT GGC ACA ATT AAC ATA TCT GGG AAT ATA Asn Leu Thr His Asn Leu Ser Gly Thr Ile Asn Ile Ser Gly Asn Ile 630	ACA ATT AAC CAA ACT AGG AGA AAG AAC ACC TGG TAT TGG CAA ACC AGC Thr 11e Asn Gln Thr Thr Arg Lys Asn Thr Ser Tyr Trp Gln Thr Ser 650	CAT GAT TCG CAC TGG AAC GTC AGT GCT CTT AAT CTA GAG ACA GGC GCA His Asp Ser His Trp Asn Val Ser Ala Leu Asn Leu Glu Thr Gly Ala 660	AAT TIT ACC TIT AIT AAA TAC AIT TCA AGC AAT AGC AAA GGC TTA ACA Asn Phe Thr Phe Ile Lys Tyr Ile Ser Ser Asn Ser Lys Gly Leu Thr 675	ACA CAG TAT AGA AGC TCT GCA GGG GTG AAT TTT AAC GGC GTA AAT GGC Thr Gln Tyr Arg Ser Ser Ala Gly Val Asn Phe Asn Gly Val Asn Gly 700	AAC ATG TCA TTC AAT CTC AAA GAA GGA GCG AAA GTT AAT TTC AAA TTA Asn Met Ser Phe Asn Leu Lys Glu Gly Ala Lys Val Asn Phe Lys Leu 710
	_			= 001		

	מחוו ככל מחוג גבל גוח חלל גיג בכי ביי ביי ביי ביי ביי ביי ביי ביי ביי	טבעב
	CCA AAC GAG AAC AIG AAC ACA AGC CCI TIA CCA AIT CGG TIT. Pro Asn Glu Asn Met Asn Thr Ser Lys Pro Leu Pro Ile Arg Phe 725	C9C7
SI	TTA GCC AAT ATC ACA GCC ACT GGT GGG GGC TCT GTT TTT TTT GAT ATA Leu Ala Asn Ile Thr Ala Thr Gly Gly Gly Ser Val Phe Phe Asp Ile 740	2613
UBSTITUTE	GCC AAC CAT TCT GGC AGA GGG GCT GAG TTA AAA ATG AGT GAA ATT Ala Asn His Ser Gly Arg Gly Ala Glu Leu Lys Met Ser Glu Ile 760	2661
SHEET (RU	AAT ATC TCT AAC GGC GCT AAT TIT ACC TTA AAT TCC CAT GIT CGC GGC Asn Ile Ser Asn Gly Ala Asn Phe Thr Leu Asn Ser His Val Arg Gly 775	2709
JLE 26)	GAT GAC GCT TITT AAA ATC AAC AAA GAC TTA ACC ATA AAT GCA ACC AAT Asp Asp Ala Phe Lys Ile Asn Lys Asp Leu Thr Ile Asn Ala Thr Asn 790	2757
•	TTC AGC CTC AGA CAG ACG AAA GAT GAT TTT TAT GAC GGG TAC Phe Ser Leu Arg Gln Thr Lys Asp Asp Phe Tyr Asp Gly Tyr 805	2805
	GCA CGC AAT GCC ATC AAT TCA ACC TAC AAC ATA TCC ATT CTG GGC GGT Ala Arg Asn Ala Ile Asn Ser Thr Tyr Asn Ile Ser Ile Leu Gly Gly 820	2853

211/235

3237	3285	3333	3381	3429	3477	3525		
GGT GAT TTA AAC ATT ACC ACT CAC GCT AAA CGC AAC CAA AGA AGC ATC Gly Asp Leu Asn Ile Thr Thr His Ala Lys Arg Asn Gln Arg Ser Ile 950	ATC Ile	AGT AAT AAT GAT GCT GAA ATC CAA ATT GGC GGC AAT ATC TCG CAA AAA Ser Asn Asn Asp Ala Glu Ile Gln Ile Gly Gly Asn Ile Ser Gln Lys 980	GAA Glu 995	ATA ACA ATC AAA AAG GGT ATT GAT GGA GAG GAC TCT AGT TCA GAT GCG Ile Thr Ile Lys Lys Gly Ile Asp Gly Glu Asp Ser Ser Ser Asp Ala 1015	ACA AGT AAT GCC AAC CTA ACT ATT AAA ACC AAA GAA TTG AAA TTG ACA Thr Ser Asn Ala Asn Leu Thr Ile Lys Thr Lys Glu Leu Lys Leu Thr 1030	GAA GAC CTA AGT ATT TCA GGT TTC AAT AAA GCA GAG ATT ACA GCC AAA Glu Asp Leu Ser Ile Ser Gly Phe Asn Lys Ala Glu Ile Thr Ala Lys 1045		
SUBSTITUTE SHEET (RULE 26)								

FIG.29k

3573	3621	3669	3717	3765	3813	3861	
GAT GGT AGA GAT TTA ACT ATT GGC AAC AGT AAT GAC GGT AAC AGC GGT. ASP Gly Arg Asp Leu Thr Ile Gly Asn Ser Asn Asp Gly Asn Ser Gly 1060	GCC GAA GCC AAA ACA GTA ACT TITT AAC AAT GTT AAA GAT TCA AAA ATC Ala Glu Ala Lys Thr Val Thr Phe Asn Asn Val Lys Asp Ser Lys Ile 1075	TCT GCT GAC GGT CAC AAT GTG ACA CTA AAT AGC AAA GTG AAA ACA TCT Ser Ala Asp Gly His Asn Val Thr Leu Asn Ser Lys Val Lys Thr Ser 1095	AGC AGC AAT GGC GGA CGT GAA AGC AAT AGC GAC AAC GAT ACC GGC TTA Ser Ser Asn Gly Gly Arg Glu Ser Asn Ser Asp Asn Asp Thr Gly Leu 1110	ACT ATT ACT GCA AAA AAT GTA GAA GTA AAC AAA GAT ATT ACT TCT CTC Thr ile Thr Ala Lys Asn Val Glu Val Asn Lys Asp ile Thr Ser Leu 1125	AAA ACA GIA AAT AIC ACC GCG TCG GAA AAG GIT ACC ACC ACA GCA GGC Lys Thr Val Asn Ile Thr Ala Ser Glu Lys Val Thr Thr Thr Ala Gly 1140	TCG ACC ATT AAC GCA ACA AAT GGC AAA GCA AGT ATT ACA ACC AAA ACA Ser Thr Ile Asn Ala Thr Asn Gly Lys Ala Ser Ile Thr Thr Lys Thr 1155	
SUBSTITUTE SHEET (RULE 26)							

3909	3957	4005	4053	4101	4149	4197
GGT GAT ATC AGC GGT ACG ATT TCC GGT AAC ACG GTA AGT GTT AGC GCG	ACT GGT GAT TTA ACC ACT AAA TCC GGC TCA AAA AITT GAA GCG AAA TCG	GGT GAG GCT AAT GTA ACA AGT GCA ACA GGT ACA ATT GGC GGT ACA ATT GLY Gly Ala Asn Val Thr Ser Ala Thr Gly Thr Ile Gly Gly Thr Ile Gly Gly Thr Ile 1205	TCC GGT AAT ACG GTA AAT GTT ACG GCA AAC GCT GGC GAT TTA ACA GTT	GGG AAT GGC GCA GAA AITT AAT GCG ACA GAA GGA GCT GCA ACC TITA ACC	GCA ACA GGG AAT ACC TIG ACT ACT GAA GCC GGT TCT AGC ATC ACT TCA	ACT AAG GGT CAG GTA GAC CTC TTG GCT CAG AAT GGT AGC ATC GCA GGA
Gly Asp ile Ser Gly Thr ile Ser Gly Asn Thr Val Ser Val Ser Ala	Thr Gly Asp Leu Thr Thr Lys Ser Gly Ser Lys Ile Glu Ala Lys Ser		Ser Gly Asn Thr Val Asn Val Thr Ala Asn Ala Gly Asp Leu Thr Val	Gly Asn Gly Ala Glu Ile Asn Ala Thr Glu Gly Ala Ala Thr Leu Thr	Ala Thr Gly Asn Thr Leu Thr Thr Glu Ala Gly Ser Ser Ile Thr Ser	Thr Lys Gly Gln Val Asp Leu Leu Ala Gln Asn Gly Ser Ile Ala Gly
1180	1190		1220	1235	1265	1270

4245	4293	4341	4389	4437	4485	4533
AGC ATT AAT GCT GCT AAT GTG ACA TTA AAT ACT ACA GGC ACC TTA ACC Ser Ile Asn Ala Asn Val Thr Leu Asn Thr Thr Gly Thr Leu Thr 1285	ACC GIG GCA GGC TCG GAT ATT AAA GCA ACC AGC GGC ACC TTG GIT AIT Thr Val Ala Gly Ser Asp Ile Lys Ala Thr Ser Gly Thr Leu Val Ile 1300		GAA GTG AAT GCA GTC AAC GCA AGC GGC TCT GGT AGT GTG ACT GCG GCA Glu Val Asn Ala Val Asn Ala Ser Gly Ser Gly Ser Val Thr Ala Ala 1335		TTA AAT ATC ATT TCG AAA GAT GGT AGA AAC ACT GTG CGC TTA AGA GGC Leu Asn Ile Ile Ser Lys Asp Gly Arg Asn Thr Val Arg Leu Arg Gly 1365	AAG GAA ATT GAG GTG AAA TAT ATC CAG CCA GGT GTA GCA AGT GTA GAA Lys Glu Ile Glu Val Lys Tyr Ile Gln Pro Gly Val Ala Ser Val Glu 1380

216/235

FIG.30A	Alignment	of HMW	proteins	
10 MNKIYRLKFSKR	LNALVAVSELAF		KGSEKPARMK	50 VRHLALKPLSAMLLS
	120 ivqfi.qennisav	130	140	150 SNGQVFLINPNGITI
	220	230		250 SDIINPTITYSIAAP
310 LSAKEGEAEIGO	JVISAQNQQAKG	GKLMITGI	340 DKVTLKTGAVI	DLSGKEGGETYLGGD
410 GNINAQGSGDIZ	420 AKTGGFVETSGH	430 DLFIKDN	440 v	
.SRKLT .KKNLT .NE.RD. .RKTD.	NL-SNGSLTLW: .IGSH.I.H .IGDSSH.I.H .IGG.SHIKDSSH.I.	SEGRSGG- .K.QR KNN .SKNKNS- NDNSS-	GVEINNDITTQ.DGSK.KESLGNSD.KGNS	40 550 GDDTRGANLTTYSGGW 5K.GQ 5ANGS 5T.GSS 5T.GSN.K. 5AT.GSVS.

217/235

FIG.30B

60	70	80	90	100	
LGVTSIPQ:	SVLASGLQGM	DVVHGTATMQ	ONDONKTIIR	NSVDAII	12-1
					12-2
1.50	150	100	100	200	
160	170			200	10 1
	NGFTASTLDI				12-1
• • • • • • • •				• • • • • •	12-2
260	270	280	290	300	
ENEAVNLG	DIFAKGGNIN	VRAATIRNQO	GKLSADSVSK	DKSGNIV	12-1
			. 	• • • • • •	12-2
360	370	380	390	400	
ERGEGKNG:	IQLAKKTSLE	KGSTINVSGF	KEKGGRAIVW	GDIALID	12-1
• • • • • • • • • • • • • • • • • • • •					12-2
	470			•	
				ILKKGTFV-NIT	
				YNAWTM	
				YSW.M	
				AR.S	
				SGN	-
				IS	
)N	
				MR.L S.MGAS	
				ANN	
				RRN.S	
	-			A.QIES	
TŐL	<u>R.D</u>	TOTACAMI.	.5111.751	.A.QID	7 1/21 1
560	570	580		590	
	LGAQGNINIT		KGS-N	OVITGO	12-1
	.DFL				12-2
T	TTL	G-S.T(GNGTEKARNA	ASSAQA.	15-1
IT	ESRL	T.EG.V	N	LT	15-2
I.IT	NSLL	TG	N	PT	Joyc-1
I.S	TFL	SNG-SV	. ADKDKARSA	ADAQ.VA.	Joyc-2

218/235

FIG.30C

.RNK.R	TIGDS.H	YKKRKNI	RSD.IQK	SG.S-	N.DD.
.TNKVN.TT	DV-YA	HRD	GN	SEKNG.	KA.S.
				STNG.	
				STNG.	
				SQNG.	
	2020 11			2	
				630	
GTITSGN	Q-KGFRFNN	VSLNGTGSGI	_QFTTKRTN	KYAITN	KFEGTLNI
.TVGEG	DA	K.	.NIISSVN	L.H	NLSI
NIG	DQ.QL.L	II.	.N.VSIQP	TSH	R.D.E.I.
A	NE.	T.	.L.NLS.PQ	KNNSLV	Y.N
A	GE.	AI.T.	L.NIDL-	GNNFQ.I.	F.N
I.NLTG	EN.TL	V.Q.	.SI.SNVG	Q.H	D.EI
ASK	D.	.T.S.VKK.I	FL.KYSQ	NNKDSNFE.1	H.R
LTG	EN.TL	N.	.SIISTAS	LSH	RLD.EI.V
NNK	DG.QL.L	A.	.N.IANQN	F.H	NIS.AI
VV	'G-Е	V.A.	.R.VGQKNIS	SNSWRENT.K.	R.D.N
AVN.	KD.	.TG.	.S.KYIE.G-	NRDSNFE.H	R.R
N.I.NÇ)EN.QLS.	M.A.	T. ANKG. $-$	н.н	D
N.1.NQ				Н.Н	
N. 1 . NÇ				н.н 720	
	690	700	710		730
DSRGS	690 DSAGTLTQ	700 PYNLNGISF	710 NKDTTFNVER	720	730 IG-INKYS
DSRGS SNSKG	690 DSAGTLTQ ELTTQYRSSA	700 PYNLNGISF .V.FVNG	710 NKDTTFNVER .M-SLKE	720 Narvnfdikap	730 IG-INKYS NE-NMNT.
DSRGS SNSKG LNNNH-GRE	690 DSAGTLTQ ELITQYRSSA TSR-YRKGG	700 PYNLNGISFI .V.FVNG GVIFRSPIG	710 NKDTTFNVER .M-SLKÐ HTNT.KQ	720 NARVNFDIKAP G.KKL	730 IG-INKYS NE-NMNT. ND-T.HAN
DSRGS SNSKG LNINNH-GRE	690 DSAGTLTÇ SLITYQYRSSA TSR-YRKGG AE.GSAP.LS	700 PYNLNGISF .V.FVNG GVIFRSPIG S.T	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNK	720 NARVNFDIKAP G.KKL GSVASFK	730 IG-INKYS NE-NMNT. ND-T.HAN TQ.N
DSRGS SNSKG LNNNH-GRE SA	690 DSAGTLTQ ELITQYRSSA TSR-YRKGG AE.GSAP.LS TNI	700 PYNLNGISF .V.FVNG GVIFRSPIG SS.T	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNK I.D.KQ	720 NARVNFDIKAP G.KKL GSVASFK KN	730 IG-INKYS NE-NMNT. ND-T.HANTQ.N
DSRGS SNSKG LNNNH-GRESAD E.NRFGPIT	690 DSAGTLIQ SLITQYRSSA TISR-YRKGG AE.GSAP.LS TNI	700 PYNLNGISF V.FVNG GVIFRSPIG S.T	710 NKDTTFNVER .M-SLKE HTNT.KQ ITNK I.D.KQ .MVLGT	720 NARVNFDIKAP G.KKL: GSVASFKKN	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.N.
DSRGS SNSKG LNNNH-GRESAD E.NRFGPTT	690 DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LS TNT -PLRSSG	700 PYNLNGISF V.V.FVNG GVIFRSPIG S.T GVFFTNG	710 NKDTTFNVER .M-SLKE HTNT.KQ ITNK I.D.KQ .MVLGT	720 NARVNFDIKAP G.KKL GSVASFKKN	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.N.
DSRGS SNSKG LNNNH-GRESAD E.NRFGPIT .NS.SRP N.ARNGDVF	690DSAGTLTQ ELITQYRSSA TSR-YRKGG AE.GSAP.LSTNI CPLRSSG -SPGP.YF	700 PYNLNGISF V.FVNG GVIFRSPIG S.T GVFFTNG RSG	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNKI.D.KQ .MVLGT .N.VAS	720 NARVNFDIKAP G.KKL GSVASFKKNGA.TSL.NL GSASP.	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK
DSRGS SNSKG LNNNH-GRESAD E.NRFGPTT .NS.SRP N.ARNGDVF	690DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LSTNIPLRSSGSPGP.YF RGRSFA	700 PYNLNGISF V.V.FVNG SVIFRSPICE S.T CVFFTNG RSG AGVIF.AKGL	710 NKDTTFNVER .M-SLKE HTNT.KQ ITNKI.D.KQ .MVLGT .NVAS TTSKK	720 NARVNFDIKAP G.KKL GSVASFKKNGA.T GSASP. GST.D.KL	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK NEKTTPNR
DSRGS SNSKG LNNNH-GRESAD E.NRFGPIT .NS.SRP N.ARNGDVFNHST	690DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LSTNI CPLRSSG -SPGP.YF RGRSFA WQDDITNRG-	700 PYNLNGISF V.FVNG GVIFRSPIG GVFFTNG RSG AGVIF.AKGL	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNK I.D.KQ .MVLGT .NVAS TTSKK EMKIGN	720 NARVNFDIKAP G.KKL GSVASFKKNGA.TSL.NL GSASP. GST.D.KL	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK NEKTTPNR XMTP.PN.
DSRGS SNSKG LNNNH-GRESAD E.NRFGPIT .NS.SRP N.ARNGDVFNHST	690DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LSTNI CPLRSSG -SPGP.YF RGRSFA WQDDITNRG-	700 PYNLNGISF V.FVNG GVIFRSPIG GVFFTNG RSG AGVIF.AKGL	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNK I.D.KQ .MVLGT .NVAS TTSKK EMKIGN	720 NARVNFDIKAP G.KKL GSVASFKKNGA.T GSASP. GST.D.KL GSTAHTS	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK NEKTTPNR XMTP.PN.
DSRGS SNSKG LNNNH-GRESAD E.NRFGPTT .NS.SRP N.ARNGDVFNHST .AS.ISSGI	690DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LSTNI TPLRSSG -SPGP.YF NSSDSRSFA NQDDITNRGSPGP.NA	700 PYNLNGISF V.FVNG GVIFRSPIG S.T GVFFTNG RSG GVIF.AKGL GVKFH.KNNT. AQSG	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNKI.D.KQ .MVLGT .NVAS TTSKK EMKIGN .GENIAQ .NVIAA	720 NARVNFDIKAP G.KKL GSVASFKKNGA.T GSASP. GST.D.KL GSTAHTS	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK NEKTTPNR WMTP.PNV-DKVIN
DSRGS SNSKG LNNNH-GRESAD E.NRFGPTT .NS.SRP N.ARNGDVFNHST .AS.ISSGI	690DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LSTNI TPLRSSG -SPGP.YF NSSDSRSFA NQDDITNRGSPGP.NA	700 PYNLNGISF V.FVNG GVIFRSPIG S.T GVFFTNG RSG GVIF.AKGL GVKFH.KNNT. AQSG	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNKI.D.KQ .MVLGT .NVAS TTSKK EMKIGN .GENIAQ .NVIAA	720 NARVNFDIKAP G.KKL GSVASFKKNGA.T GSASP. GST.D.KL GSTAHTS .SSANP.	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK NEKTTPNR WMTP.PNV-DKVIN
DSRGS SNSKG LNNNH-GRESAD E.NRFGPTT .NS.SRP N.ARNGDVFNHST .AS.ISSGISGSAS	690DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LSTNI PLRSSG -SPGP.YF RGRSFA NQDDITNRGSPGP.NA	700 PYNLNGISF V.V.FVNG GVIFRSPIG S.T GVFFTNG RSG GVIF.AKGL GVKFH.KNNT. AQSG	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNKI.D.KQ .MVLGT .NVAS TTSKK EMKIGN .GENIAQ .NVIAA	720 NARVNFDIKAP G.KKL GSVASFKKNGA.T GSASP. GST.D.KL GSTAHTS .SSANP.	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK NEKTTPNR WMTP.PNV-DKVIN
DSRGS SNSKG LNNNH-GRESAD E.NRFGPITI .NS.SRP N.ARNGDVFNHST .AS.ISSGNSGSASNRSVA	690DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LSTNI -PLRSSG -SPGP.YF RGRSFA NQDDITNRG SPGP.NF -LNSGSRSFA	700 PYNLNGISF V.Y.FVNG GVIFRSPIG SS.T GVFFTNG RSG AGVIF.AKGL AGVKFH.KNN AQSG AGVKFY.KNN	710 NKDTTFNVER .M-SLKE HTNT.KQ ITNKI.D.KQ .MVLGT .N.VAS TTSKK EMKIGN .GENIAQ .NVIAA	720 NARVNFDIKAP G.KKL GSVASFK KN .GA.T GSASP. GST.D.KL GSTAHTS SSANP.	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK NEKTTPNR TWMTP.PNV-DKVIN IDNTSNNKP
DSRGS SNSKG LNNNH-GRESAD E.NRFGPTT .NS.SRP N.ARNGDVFNHST .AS.ISSGISGSASNRSVA 75 RFKTSGSTI	690DSAGTLTQ ELTTQYRSSA TSR-YRKGG AE.GSAP.LSTNI TPLRSSG -SPGP.YF RGRSFA NQDDITNRGSPGP.NA -LNSGSRSFA	700 PYNLNGISF V.FVNG GVIFRSPIG S.T GVFFTNG RSG GVFF.AKGL GVKFH.KNN AQSG AGVKFY.KNN B00 LTLNATGGNI	710 NKDTTFNVER .M-SLKE HTNT.KQ TTNKI.D.KQ .MVLGT .NVAS TTSKK EMKIGN .GENIAQ .NVIAA EMKIGD	720 NARVNFDIKAP G.KKL GSVASFKKNGA.T GSASP. GST.D.KL GSTAHTS SSANP. 0N.E.KL.SN	730 IG-INKYS NE-NMNT. ND-T.HANTQ.NV.NNR NE-NT.NV-S.VHD NSGYNSQK NEKTTPNR WMTP.PNV-DKVIN IDNTSNNKP 830 AKKNITFEG

FIG.30D

T.NSYLT.SG.VQ.NDLT I.G.T.EFLSSD-SVG.NG.KGRSSASAQ.IA. T.EFLSGNLTA. I.T.NTYLGG-SVAGNEKGRQVSESKA. T.KSYLTGDKPGLSNLTAK T.MFLSDN-N.TDLTA.	LCDC2-1 LCDC2-2 PMH1-1 PMH1-2 K1-1 K21-1
640 650 660 670 680 SGKVNISMVLPKNESG-YDKFKGRTYWNLTSLNVSESGEFNLTI .NIT.NQTTRTWQTSHDSHVSALETGAN.TFIKYR.HVNQTTLFW.VSDESVSH.T.K.KSA.SF.KFA .SIP.NAT.NW.SRYI.H.ADSN IKWDS.RV.HGSK T.N.T.NQTA.ATTAWN.SYDSVSTQKNSS.TFIKRTD.L.QARQENWN-RRHSSHV.RTNSYL.I N.T.NQTTQQ.IEW.ASSDSV.F.LR.DSK.TFIKYVV.T.NQTTKAKAWNTSYDSVST.TL.NDAK.TFIKYVV.DVSGTKWH-TRINV.TASGSSSD.L.QARQENWN-RRHWSHV.RNSYVV.NQTT.HIA-PWNASADSV.T.TLGNNAQ.TFIKFV	12-1 12-2 15-1 15-2 Joyc-1 Joyc-2 LCDC2-1 LCDC2-2 PMH1-1 PMH1-2 K1-1 K21-1
740 750 760 770 780 SLNYASFNENISVSGGGSVDFTLLASSSNVQTPGVVINSKYFNVSTGSSL KPLPIR.LA. TAT F.DIY.NHGR. AELKMSEI . I . N. ANF -QLPIQ S D . K. L. CITSNYGRS . G. GMSSI D . N. N L N. T. R. N Q I HL. A. K N	
840 850 860 870 880 GNITFGSRKAVTEIEGNVTINNNANVTLIGSDFDNHQKPLTIKK-DVI	

FIG.30E

T.NS.IRGQEA.N.SISFFE.G.YSD.FNGNGFNHDA.KSTHSIL.
E.TVL.NNSIEVT.
E.RAEVL.NNSIVA.
TLQSHVRKDSA.I.SIS.FE.SPDSFT.KYP.R-A.SSTIS.
SER.A.TSS.NA.INLQSLN
SMTAQARDRNA.E.TVISNS.LSII.QNDGFDNNQKAN-A.NS.Y.V.IQ.
TLNSHVRKYNA.E.NINS.FN.R.TSDSFRN.YRNN-A.NSTHSIL.
NL.AN.A.L.KNNEIKSR.QVEQI
K.SE.HAA.T.KN.I.S.N.A.ISNLKSLI.N
SINSHVRGNNA.E.KIIS.FN.K.TKDKFDNSYEKN-A.FSTH.L.IL.
890 900 910 920 930
INSCNLTAGGNIVNIAGNLTVESNANFKAITNFTFNVGGLFDNKGNSNISIAKGGA
LVN.S.SLT.ENAD.KISES.TGK.RD.L.IT.N.T.N.TAE.N.TQ.VV
.ESLI.ASANS.KEKGE.QDNL.IT.T.I.N.D.K.N.SQ.VV
NRNGLN
V-ELT.SVAD.KSILND.TGE.SENL.IT.N.T.N.TAD.N.KQ.VV
KVT.SAIEKGS.K.L.NP.YSSQ.K
VEE.RLV.ASANNS.K.G.KEDNL.IT.T.T.N.T.I.DVKA.
VESLI.ENANSI.KE.IGK.KDSL.IT.N.T.N.TAE.N.SQ.VV
TAND.DTIK.KD.AQG.K.NGS.KNNL.IT.T.T.N.T.I.D.TQ.VV
NKVT.SAIEKGS.K.L.NP.YSSQ.K
SVGNII.SNAH.DSIAES.K.QGKNNL.IT.T.T.N.TAD.N.KQ.VV
990 1000 1010 1020 1030
TEMQIGGDVSQKEGNLTISSDKINITKQITIKAGVDGENSDSDATNINANLTIKTKE
A.INIK.ID.SS
A.INI
A.INIKNSTKSQ
A.ININK.NK.DSTA
A.INIKKECGSPAST
A.ININKNK.DSTA
.INI
A.IENI
A.IENI
.INIVERN.DNEATS
A.INI

FIG.30F

V.L.GQNSSSS.TIEKAEANNAP.Q.NIRDRVLG-SL	12-2
V.L.GQDSSST.TIN.SQARAYNGNGRNQLGN.S	15-1
KIKE.T.AND.KNGV	15-2
KIA	Joyc-1
VSL.GQNSSSD.KIKSSTKAHNSPRDFASRTLGNLN	Joyc-2
LAAD.KPIKI.VKEGRSANYG.DKSA.S.R-GN.T	LCDC2-1
V.L.GQNSSST.T.S.N.GAQAHNGNDRNKFGN.S	LCDC2-2
V.L.GQNSSSS.MII.KRAEADNSH.SDNV.DRNLGNLT	PMH1-1
L.Q.P.K.D.VKQGT.A.RSANGAL.VNGN	PMH1-2
LAAD.KPIKI.VKEGRSANYG.DKSA.SRGN.T	K1-1
V.L.GENSSSN.KINSKQAHAGTS.LDK.ER.LTLGN	K21-1
~	
940 950 960 970 980	
RFK-DIDNSKNLSITINSSSTYRTIISGNITNKNGDLNITNEGSD	12-1
KLG-NVT.DGD.NHAKRNQ.SG.D.IK.SDSNN.	12-2
KLG-NVT.DGD.NHAKHNQ.SG.D.IK.SDSNKN	15-1
KN.TSS.NDTEAIDNKGN	15-2
KN.TSS.NDTEAIDNKGN	Joyc-1
NIQGN.T.KGG.NAQNNQKNEGKDSNNN	Joyc-2
HN.T.S.ND.AESDNKNN	LCDC2-1
KLG-N.T.DGNAKNGQKSV.NNK.AN.N.	LCDC2-2
SLGT.DGK.NHAK.GQKS <u>R.D</u> .IQ.NDNN.N	PMH1-1
NLG-NVT.DGK.NHAK.GQKS <u>R.D</u> .IQ.NDNN.N	PMH1-2
IE.TGS.NK.D.NHHKRKN.DN	K1-1
KLQGT.NGNA.VNQKNKDIKAN	K21-1
1040 1050 1050 1050 1000	
1040 1050 1060 1070 1080	10 1
LKLTQDLNISGFNKAEITAKDGSDLTIGNIN-SADGINAKKVTFN	12-1
ESRS.DGNS.AET	12-2
	15-1
K.VS.NSDD.GNST	15-2
.QGDE.AISDNNNA	Joyc-1
.EG	Joyc-2 LCDC2-1
.QGDE.AISDNNNA ENDVENNN.ING.NAT	LCDC2-2
T. DN	PMH1-1
.TF.DN	PMH1-1
. II	FLIII_{-7}

FIG.30G

	1090	•			
N N K N K		LHSKVET-SGS .NKSNK.L.DM .NNK .NNK .NNK .NNK .NNK .NNK .NNK .NNK .NNK	GGR.SNI VIDG.G.G. VIDGG GDA.SNNGD: DSSADN GDA.SNNGD: GDA.SNNGD: DGS.GNGD DGS.GNGD	OTTATTD .TSD .TSD .TSD	.EKDE.KE.KDDV IDS
		NN 1210	.DGS.GNGD	.NISI)S
			NSGALTTLAGS 		
		A T.QNA			
		I			
A					
NATT	GDANITTQIC		SSVTL1ATGQTT 	_AVGN1SGDTV 	TTTADKGKLIT
	DVNITTST(GSINGKIESKS	GSVTLTATEKTI	LTVGNVSGNIV	TVTANRGALT

FIG.30H

NN.NDNSDGN.D TDNN.NIKAS.NSQI.	
1140 1150 1160 1170 1 HKAVSISATSGEITTKTGTTINATTGNVEITAQTGSILGGIES L.T.N.T.SEK-V. TA.S. N.KAS. TK. D.S N.T.N.T.SEK-L. ADA. V. K. D.K.EVKT.NV. AN.G. A. H. QT.N.T.SEN-V. A. I.S. V. K. D.KTIN. T.NV. ES. A. S. V. K. D.ST.N.T.SEN-V. A. S. V. K. D.K.KVT.N.T.SER-D. AD KL. V.SD KT.N. SE.G. A. S. V. KT.N. SE.G. A. S. V. KT.N.T.SEN-V. A. S. V. K.D.KT.N.T.SEN-V. A. S. V. KT.N.T.SEN-V. A. S. V. K.D.KT.N.T.SEN-V. A. S. V. K.D.K.	T 12-2 T 15-1 KP. 15-2 N Joyc-1 Joyc-2 T LCDC2-1 N LCDC2-2 PMH1-1 PMH1-2 N K1-1
EVKSASGNVNITASGNTLNVSNITGQNVTVTANSGAITTTEG EVKSASGNVNITASGNTLNVSNITGQNVTVTANSGAITTTEG	12-2 15-1 15-2 STI Joyc-1 Joyc-2 LCDC2-1 LCDC2-2 PMH1-1 PMH1-2
1220 1230 1240 1250	12-1 12-2 15-1 15-2 Joyc-1 Joyc-2

FIG.301

224/235

NATTGDANITTQTGNINGKVESSSGSVTLIATGQTLAVGNISGDTVTITADKGKLT -----TDLTTVKGGAKINATEGTATLTASSGKLT TTOADSKIEATEGEANVTSKTSIIGGTISGGTVEVTATEGLTTOAGSTITGTESVT 1280 1290 1260 1270 TTOSNSKIKATTGEANVTSATGTIGGTISGNTVNVTANAGDLTVGNGAEINATEGA ..K.G...E.KS...... -----...TLNVS----.VSGN. I.K.G.E....A..V......D.....T.....EDA.K.D..G.. _____ -----KG..K.....T ----.D.S....K..S...ST.....RKA.T.S.... -----.K..S...ST.....RKA.T.S.... 1370 1380 1390 1400 1360 LNITGILTIVKGSNINATSGILVINAKDAFLNGAALGNHIVVNATNANGSGSVIATA..D.K......K...D.S.DS.E...V..S.....T.AE....A.....K....S.D......S....T.VA.....S...N........E.S.....A....K.D.T.S.DR......S.....T.A

225/235

FIG.30J

TLAGSTIN. NG. E. EVT.K.S.T.AG. TOTSSKIN. K. S. N. S.S.G TEANSAIS.ANG. A. S. K.S.T.SSG.	LCDC2-1 LCDC2-2 PMH1-1 PMH1-2 K1-1 K21-1
TSSQSGNIGGMISGGKVEVSATKDL	12-1 12-2 15-1 15-2 Joyc-1 Joyc-2 LCDC2-1 LCDC2-2 PMH1-1 PMH1-2 K1-1 K21-1
1310 1320 1330 1340 1350 ATLTTSSGKLTTEASSHITSAKGQVNLSAQDGSVAGSINAANVT . ATGNT . G.S. T D.L . N. I	12-1 12-2 15-1 15-2 Joyc-1 Joyc-2 LCDC2-1 LCDC2-2 PMH1-1 PMH1-2 K1-1 K21-1
1410 1420 1430 1440 1450 TSSRVNITGDLITINGLNIISKNGINTVLLKGVKIDVKYIQPGIS. N.V. D.R. R.R.KE.EVNS.VRVTE.EV	12-1 12-2 15-1

226/235

FIG.30K

AS	.N	E.S	
			ET.K
			ET.K
			.EST.K
			EVSN.T.K
D.	.IVEQ.	.1	
1460	1470 14	1490	1500
ASVDEVIEAKRILEKV	KOLSDEEREALA	KLGVSAVRFIE	PNNTITVDTQNEFATRP
EV	T	V.	NT
EV	T		NT
• • • • • • • • • • • • • •		A.	AINT
YA			
ANA			
EV			
EV			
			AINT
			AINT
			NT
		• • • • • • • • • • •	

WO 00/20609

PCT/CA99/00938

227/235

FIG.30L

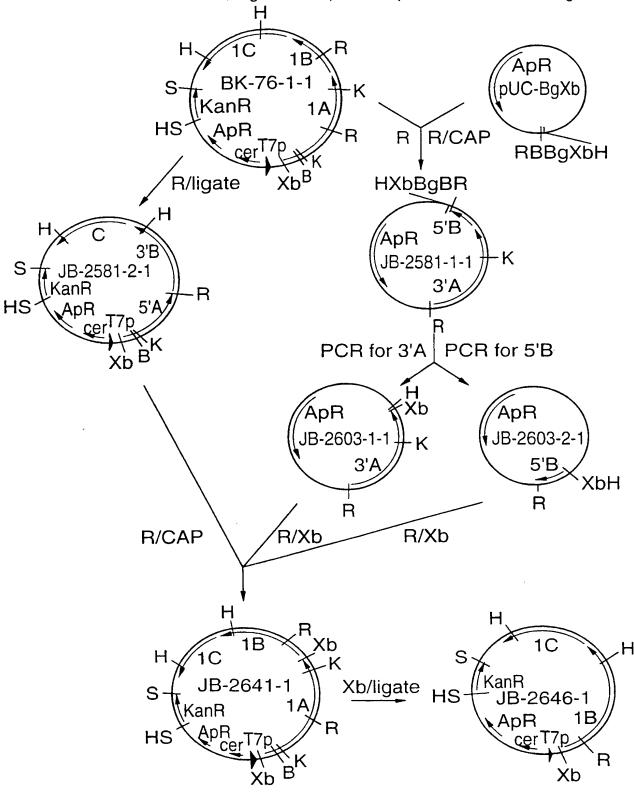
	15-2 Joyc-1 Joyc-2 LCDC2-1 LCDC2-2 PMH1-1 PMH1-2 K1-1 K21-1
1510	12-1 12-2 15-1 15-2 Joyc-1 Joyc-2 LCDC2-1 LCDC2-2 PMH1-1 PMH1-2 K1-1

228 /235

um/2.				
ies were himul or l	SEO ID NO 74	75	76 77 78	79 80 81
used to determine whether PCR amplified hmwA genes were hmw1 or hmw2.	5672.SL	78.979	5742.SL	5743.SL
whether PCR a	3,	3,		
rmine		<i>r</i> >	, ,	5,
Oligonucleotides used to dete	TCTTTTGCTGTGGCTGATGCCCCTA	CACTGATAGGITGCTCATATTCGCC	V G V H K N GGITGATGITCATAAAAATAT CCAACTACAAGTATTITITATA	G G S L T I N S GGCGGAAGTTTAACTATTAACTC CCGCCTTCAAATTGATAATTGAG
Oliç	5,	5,	3,	3,

229/235

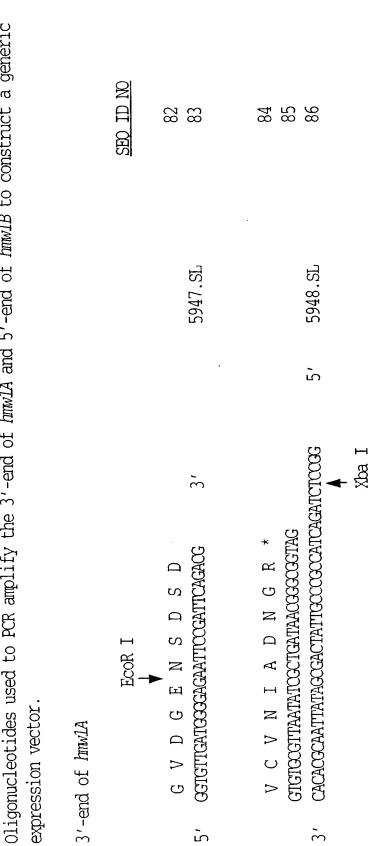
Construction of JB-2646-1, a generic expression plasmid for hmwABC genes

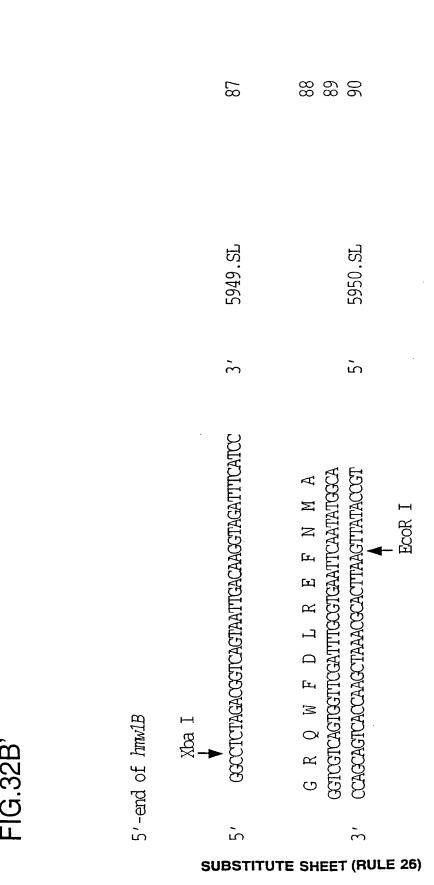


pT7 hmw1BC(12)/cer/KanR

FIG.32A

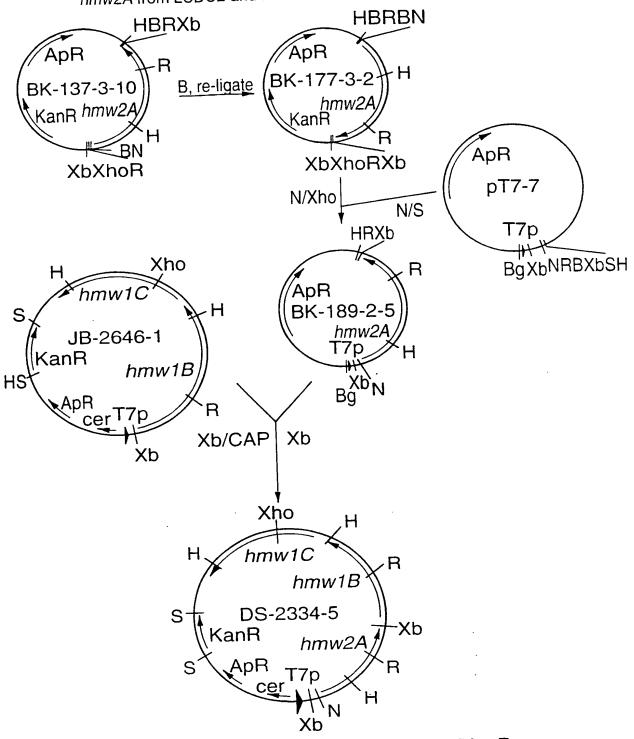
Oligonucleotides used to PCR amplify the 3'-end of hmwlA and 5'-end of hmwlB to construct a generic





232/235

Construction of DS-2334-5, a chimeric expression plasmid with hmw2A from LCDC2 and hmw1BC from strain 12



pT7hmw2A (LCDC2)/hmwBC (12)/cer/KanR

FIG.33A

Oligonucleotides used to PCR amplify the LCDC2 hmw2A gene for expression

sense

233/235

Bamii Nde I What is a second control of the	3,	5972.SL	SEO ID NO 91 92
antisense A A V C T N V A D D G Q Q * 5' GCAGCAGTATGTACCAATGTTGCTGACGATGGACAGCAGTAGT 3' CGTCGTCATACATGGTTACAACGACTGCTACCTGTCGTCATCAGATCTG XBA I	ý, ý,	5973.SL	93 94 95

234/235

Construction of DS-2400-13, a pBR T7 hmwA/T7 hmwABC/cer/KanRplasmid

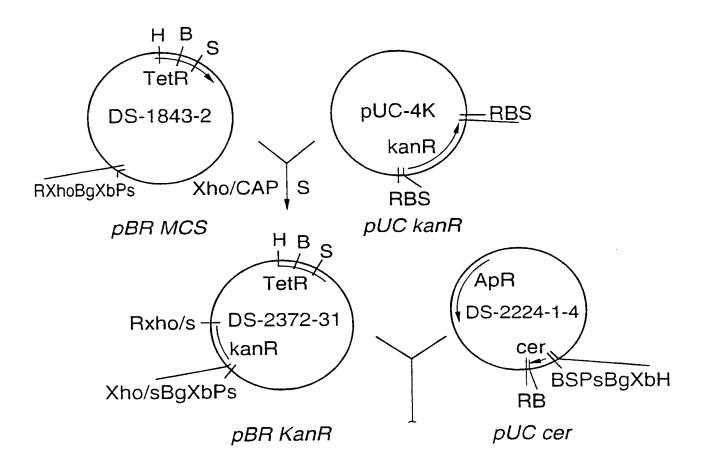
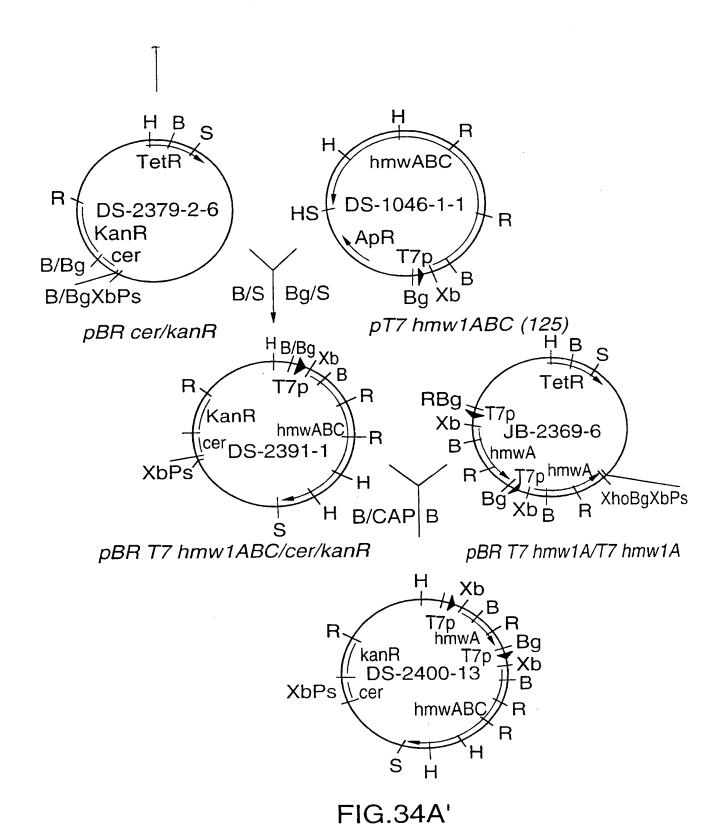


FIG.34A



SUBSTITUTE SHEET (RULE 26)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:
☐ BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
BLURRED OR ILLEGIBLE TEXT OR DRAWING
SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

IMAGES ARE BEST AVAILABLE COPY.

OTHER: ___

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.